

#4
BT
12-12-01

AO

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re patent application of

Kamiya et al.

Serial No.: 09/966,099

Group Art Unit: 2661

Filing Date: October 1, 2001

Examiner: Unknown

For: TWO-DIMENSIONAL PIPELINED SCHEDULING TECHNIQUE

Assistant Commissioner of Patents
Washington, D.C. 20231

RECEIVED
DEC 11 2001
Technology Center 2600

SUBMISSION OF PRIORITY DOCUMENT

Sir:

Submitted herewith is a certified copy of Japanese Application Number 2000-302551
filed on October 2, 2000, upon which application the claim for priority is based.

Respectfully submitted,

Sean M. McGinn

Registration No. 34,386

Date: 12/7/01

McGinn & Gibb, PLLC

Intellectual Property Law

8321 Old Courthouse Road, Suite 200

Vienna, Virginia 22182-3817

(703) 761-4100

Customer No. 21254

FOS-562
VS



日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日
Date of Application:

2000年10月 2日

出 願 番 号
Application Number:

特願2000-302551

出 願 人
Applicant(s):

日本電気株式会社

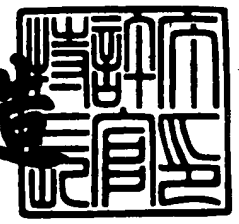
RECEIVED
DEC 11 2001
Technology Center 2600

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年 8月17日

特許庁長官
Commissioner,
Japan Patent Office

及 川 耕 造



【書類名】 特許願

【整理番号】 49220160PY

【あて先】 特許庁長官殿

【国際特許分類】 H04L 12/28

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 神谷 聡史

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 尾崎 博一

【特許出願人】

 【識別番号】 000004237

 【氏名又は名称】 日本電気株式会社

【代理人】

 【識別番号】 100083987

 【弁理士】

 【氏名又は名称】 山内 梅雄

【手数料の表示】

 【予納台帳番号】 016252

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

 【包括委任状番号】 9006535

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 パケット交換装置

【特許請求の範囲】

【請求項 1】 複数の入力ポートのそれぞれに用意され、入力されるそれぞれのパケットを宛先の出力ポートに対応した格納場所に順に格納するバッファメモリと、

それぞれの入力ポートのバッファメモリに格納されたパケットについての転送要求の有無を表わした情報を、予め全入力ポートと全出力ポートを重複しないように等しい数ずつ選択してなる入力ポートと出力ポートの組み合わせからなる $M \times M$ のマトリックス状に配置された複数のグループのうちの該当する入力ポートと出力ポートに対応した 1 つずつのグループに割り振るグループ別割振り手段と

このグループ別割振り手段によって割り振られたそれぞれのグループの中で、前記 $M \times M$ のマトリックス上で行および列方向に競合が生じない M 個のグループを選択したときのこれらのグループを前記 $M \times M$ のマトリックスの行方向と列方向の 2 方向にパイプライン処理して前記固定長セルの転送要求を表わした情報の予約をグループ単位で行う予約処理手段と、

この予約処理手段の予約処理結果に対応させて前記バッファメモリに格納されたパケットの入出力を設定して転送要求のあった前記パケットの転送を行うスイッチ

とを具備することを特徴とするパケット交換装置。

【請求項 2】 複数の入力ポートに入力されるそれぞれのパケットを所定の固定長セルに分解する固定長セル分解手段と、

それぞれの入力ポートに対応して配置され、この固定長セル分解手段によって分解されたそれぞれのセルを宛先の出力ポートに対応した格納場所に順に格納する論理キューと、

それぞれの入力ポートの論理キューに格納された固定長セルについての転送要求の有無を表わした情報を、予め全入力ポートと全出力ポートを重複しないように等しい数ずつ選択してなる入力ポートと出力ポートの組み合わせからなる $M \times$

Mのマトリックス状に配置された複数のグループのうちの該当する入力ポートと出力ポートに対応した1つずつのグループに割り振るグループ別割振り手段と、

このグループ別割振り手段によって割り振られたそれぞれのグループの中で、前記M×Mのマトリックス上で行および列方向に競合が生じないM個のグループを選択したときのこれらのグループを前記M×Mのマトリックスの行方向と列方向の2方向にパイプライン処理して前記固定長セルの転送要求を表わした情報の予約をグループ単位で行う予約処理手段と、

この予約処理手段の予約処理結果に対応させて前記論理キューに格納された固定長セルの入出力を設定して転送要求のあった前記固定長セルの転送を行うスイッチと、

このスイッチによってそれぞれの出力側に得られた固定長セルをパケットに組み立てるパケット組立手段

とを具備することを特徴とするパケット交換装置。

【請求項3】 前記予約処理手段は、1つのグループの予約から次のグループへの予約結果の転送までを1つのタイムスロットで行うことを特徴とする請求項1または請求項2記載のパケット交換装置。

【請求項4】 予め全入力ポートと全出力ポートを重複しないように等しい数ずつ選択してなる入力ポートと出力ポートの組み合わせは、ダイアゴナル・サービス・パターンであり、そのM個の組み合わせからなるダイアゴナル・モジュール群を用いてパイプライン処理が行われることを特徴とする請求項1または請求項2記載のパケット交換装置。

【請求項5】 前記予約処理手段は、予め定めた未来の所定のタイムスロットに対して、前記ダイアゴナル・サービス・パターンに該当するグループから該当する出力ポートに対して順次転送要求予約を行うことを特徴とする請求項4記載のパケット交換装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明はパケットスイッチを備えたパケット交換装置に係わり、特に高速でス

イッチ動作を行うことのできるパケット交換装置に関する。

【0002】

【従来の技術】

インターネットが爆発的に普及している。これに伴って公衆通信インフラストラクチャの急激な変革が求められている。また、インターネットの枠組み自体を通信インフラストラクチャとして活用しようとする機運が高まっている。インターネットがこのような役割を果たすためには、サービスノードとしてのルータにデータ転送処理の高速化および高機能化が求められている。

【0003】

現在の高速ルータでは、ハードウェアによるIP (internet protocol) アドレス検索と、セルフルーティングによる高速スイッチファブリックを用いたデータ転送処理が適用されている。

【0004】

高速大容量ルータのスイッチを実現するために、VOQ (Virtual Output Queuing) を使用した入力バッファスイッチが広く使用されている。従来のFIFO (first-in first-out: 先入れ先出し) 式の入力バッファ型スイッチではスループット低下の問題があったが、これを解決したノンブロッキングスイッチとして高速化が可能となっている。

【0005】

図15は、VOQを使用した入力バッファ型スイッチの構成を表わしたものである。このスイッチ100は、第1～第Nの入力インタフェース部101₁～101_Nと、第1～第Nの出力インタフェース部102₁～102_Nと、これらの間に配置されたスイッチ部103によって構成されている。

【0006】

第1～第Nの出力インタフェース部102₁～102_Nのそれぞれは、宛先解決エンジン部 (FE) 104、パケット生成分解部 (PAD) 105およびVOQ 106から構成されている。ここで、宛先解決エンジン部104は図示しない伝送路から送られてきたパケット (図示せず) の宛先の解決を行う。パケット生成分解部105は、このパケットスイッチの転送単位の固定長セルに分解する。分

解されたそれぞれのセルは、宛先に応じてVOQ 106内の出力インタフェース部 $102_1 \sim 102_N$ に対応したバッファに格納される。

【0007】

スイッチ部103は、クロスポイントスイッチ108とスイッチスケジューラ109とによって構成されている。ここでスイッチスケジューラ109は、第1～第Nの入力インタフェース部 $101_1 \sim 101_N$ のVOQ 106から得られた情報を基にしてこれら第1～第Nの入力インタフェース部 $101_1 \sim 101_N$ のセル転送要求を調停して、それぞれのVOQ 106にセル転送許可を与える。また、クロスポイントスイッチ108を制御して、VOQ 106から出力される調停後のセルを出力インタフェース部 $102_1 \sim 102_N$ のうちの対応する宛先へスイッチングさせるようになっている。

【0008】

第1～第Nの出力インタフェース部 $102_1 \sim 102_N$ は、VIQ (Virtual Input Queuing) 111とパケット組立部(PAD) 112から構成されている。スイッチ部103から送られてきたセルは、第1～第Nの出力インタフェース部 $102_1 \sim 102_N$ における対応する出力インタフェース部のVIQ 111が受け取り、パケット組立部112で元のパケットに組み立てる。そして図示しない伝送路に出力することになる。

【0009】

さて、最近のパケット交換装置では、図15に示した入力バッファ型スイッチのようなN入力およびN出力(Nは2以上の整数)のクロスポイントスイッチ(クロスポイントスイッチ108参照)が使用されることが多い。この場合には、入力部にN個の出力それぞれに対応した論理的待ち行列(論理キュー)(VOQ 106参照)を備えると共に、各論理キューから発せられるパケット転送要求を調停して、クロスポイントスイッチの接続の制御を行うスケジューラ(スイッチスケジューラ109参照)が必要となる。

【0010】

複数の待ち行列を有するパケット交換装置でこのようなスケジューリングを行う技術は、特開平6-233337号公報や特開平9-326828号公報に開

示がある。このうち特開平 6 - 2 3 3 3 3 7 号公報に記載された技術では、入力ポートを表わす各行と、出力ポートを表わす各列とを備えた要求マトリックスを使用している。マトリックスの所与の行および列のビットが、対応する入力ポートから出力ポートへの接続要求を表わしているものとする。この技術では、マトリックス上で競合が生じない N 個の要素を選択して、これをダイアゴナル・サービス・パターンとしている。

【 0 0 1 1 】

ダイアゴナル・サービス・パタンの典型的な例は対角要素の集合である。マトリックス全体が重複なくすべて覆われるように N 個ずつの異なるダイアゴナル・サービス・パターンを決定し、スケジューリングでは図 1 5 で説明した分解後のパケットスイッチ転送単位の固定長セルの処理単位としての 1 タイムスロット内にダイアゴナル・サービス・パターンを順次、接続要求マトリックス上で適用して要求を処理していく。開始のダイアゴナル・サービス・パターンは、タイムスロットごとにラウンドロビンで適用するようにしている。

【 0 0 1 2 】

一方、特開平 9 - 3 2 6 8 2 8 号公報に記載されたスケジューリングの技術では、既に説明した特開平 6 - 2 3 3 3 3 7 号公報に記載された技術と同様にそれぞれの入力からそれぞれの出力への転送要求をマトリックス上に並べ、対角線方向に検索と割り当てを行うようにしている。この技術では、タイムスロットごとに入力と出力をそれぞれランダムに並び替えるようにしている。これにより各入力側のパケットの処理の公平性が図られている。

【 0 0 1 3 】

ところが、以上簡単に説明した従来のスケジューリングの手法では、入力ポートおよび出力ポートの数が増大すると、1 タイムスロット内に行うべきデータの処理量がポート数の二乗に比例して増大し、このために高速なスケジューリングが困難になるという問題があった。特に最近ではインターネットにおけるデータの伝送量が増大するだけでなくデータ伝送の高速化が強く求められており、いまままでと同様の入出力ポートの数であったとしても、より高速のスケジューリングが求められている。そこで、この問題を解決するためにスケジューリングをパイ

プライン化する技術が登場している。

【0014】

図16は、この技術を使用したスイッチ部およびその周辺の構成を示したものである。なお、この図16に示した部分は、図15におけるスイッチ部103とその周辺に対応する部分である。スイッチ部121は、スイッチ素子としての $N \times N$ のクロスポイントスイッチ122と、スイッチスケジューラ123によって構成されている。このスイッチ部121には N 本の入力ライン 124_1 、 124_2 、…… 124_N と接続された入力ポートと、同じく N 本の出力ライン 125_1 、 125_2 、…… 125_N と接続された出力ポートが配置されている。各入力ポートは図15で示したVOQ106と同様のVOQ126を備えている。それぞれの入力ポートのVOQ126にはスイッチスケジューラ123内の分散スケジューリング・モジュール 127_1 、 127_2 、…… 127_N が対応して配置されている。

【0015】

スイッチスケジューラ123はタイムスロット単位に入力ポートから出力ポートごとの接続要求情報(Request)を受信して、これを基にしてこれらの間の接続許可情報(Grant)を決定する。スイッチスケジューラ123はまた接続許可情報を基にして入力ポートと出力ポートの接続情報(MSEL)を生成して、クロスポイントスイッチ122に通知して、クロスポイントスイッチ122の入出力の接続を設定する。また、スイッチスケジューラ123は接続許可情報を元にして、各入力ポートがどの出力ポートへのデータ転送を許可されているかを示す転送許可情報(DSTMSG)を作成して、各入力ポートに対して転送許可情報を通知する。入力ポートは転送許可情報にしたがって、データをクロスポイントスイッチ122へ出力して、出力ポート側がデータを受信することでスイッチングが完了するようになっている。

【0016】

このようなスイッチスケジューラ123の機能は、 $N \times N$ の接続要求情報から $N \times N$ の接続許可情報を生成することである。接続許可情報生成するにあたって、各分散スケジューリング・モジュール 127_1 、 127_2 、…… 127_N は、個々の入力ポートに対する出力ポートへの接続の可否を決定する。

【0017】

ところで、1999年にラウンド・ロビン・グリーディ・スケジューリング (RRGS) アルゴリズムが発表された。スイッチスケジューラ123はこのRRGSアルゴリズムを用いている。RRGSアルゴリズムを用いたスイッチスケジューラ123は、分散スケジューリング・モジュール 127_1 、 127_2 、…… 127_N がリング状に接続されている。そして、これらの隣接したモジュール間でメッセージの受け渡しが行われる。RRGSアルゴリズムでは、各分散スケジューリング・モジュール 127_1 、 127_2 、…… 127_N が、対象となるタイムスロットの予約（接続許可決定）を行い、その結果情報を次の分散スケジューリング・モジュールに渡すようにしている。

【0018】

RRGSアルゴリズムを使用したスケジューラでは、分散スケジューリング・モジュールがリング状に接続され、隣接した分散スケジューリング・モジュール間でメッセージ受渡しを行う。RRGSアルゴリズムでは、各分散スケジューリング・モジュールが対象となるタイムスロットの予約（接続許可決定）を行い、結果の情報を次の分散スケジューリング・モジュールに渡す。メッセージ受渡し速度要求条件を緩和するために、RRGSはパイプライン機能を導入している。あるタイムスロットの予約過程は、各分散スケジューリング・モジュール間でメッセージ受渡しが一周することにより完了する。また、RRGSアルゴリズムでは、N個の分散スケジューリング・モジュールが、現在のスロットの少なくともNスロット先のタイムスロットに関して予約する。更に、RRGSアルゴリズムは、N個のタイムスロットに対する予約過程を1タイムスロットずつ位相をずらしながら同時に進行させていく。

【0019】

一方、RRGSアルゴリズムの変形として、複数のタイムスロットに対する予約過程をそれぞれ異なる分散スケジューリング・モジュールから同時に開始し、進行させ、同時に終了させるアルゴリズムも考えられる。これをフレーム化RRGSと呼ぶこととする。

【0020】

図17はRRGSおよびフレーム化RRGSを使用する分散型スケジューラの構成を示したものである。ここでは一例としてポート数Nが“4”の場合を示している。スイッチスケジューラ141は分散型スケジューリングのための第1～第4のインプットモジュール (Input Module; IM) 151₁～151₄から構成される。インプットモジュール151₁～151₄には、フレームの先頭を示すフレームパルス (FP) 152が入力される。各インプットモジュール151₁～151₄はフレームパルス152に同期して動作する。

【0021】

各インプットモジュール151₁～151₄にはそれぞれのインプットモジュールを識別するための物理番号155₁～155₄が設定される。第1～第4の各入力ポートからそれぞれ対応する接続要求情報156₁～156₄がインプットモジュール151₁～151₄に入力されると、これらのインプットモジュール151₁～151₄は接続要求の調停を行い、その結果として予約（接続許可）を決定し、接続許可情報157₁～157₄を出力するようになっている。

【0022】

RRGSおよびフレーム化RRGSでは、隣接分散スケジューリング・モジュール間で接続許可情報から入力ポート情報を縮退させた情報（入力ポート情報を参照して作成した情報）としての「出力ポート予約済情報」を受け渡すことにより、出力ポートに対する接続要求の競合を回避している。例えば、第3のインプットモジュール151₃は、前段の第2のインプットモジュール151₂から出力ポート予約済情報161₂を出力ポート予約済情報162₃として受信して接続要求の調停に使用する。接続許可情報決定後、出力ポート予約済情報161₃を次段の第4のインプットモジュール151₄に通知するようになっている。

【0023】

図18はフレーム化RRGSの一般的な動作例を示したものである。図17と共にフレーム化RRGSの動作を説明する。ここでは、4×4の入出力ポートを有するスイッチの動作となる。

【0024】

図17に示した出力ポート予約済み情報161の転送方向は、インプットモジ

ジュール151の番号が増加する方向であり、この例では第4のインプットモジュール151₄まで転送したら再び第1のインプットモジュール151₁に戻ってここから第2のインプットモジュール151₂、第3のインプットモジュール151₃、……という順序で転送が行われる。

【0025】

接続要求情報156等の各情報の極性は次のように定義される。まず、第1～第4の入力ポートの接続要求156₁～156₄は、信号“1”が要求ありを示し、信号“0”が要求なしを示す。第1～第4の入力ポートの接続許可情報（予約情報）157₁～157₄は、信号“1”が許可（予約済み）を示し、信号“0”が禁止（未予約）を示す。出力ポート予約済み情報161₁～161₄は、信号“1”が予約済みを示し、信号“0”が未予約を示す。

【0026】

またこの例では4タイムスロットが1フレームを構成しており、フレームパルス152は4タイムスロット周期で入力される。図18では、第1～第4の各タイムスロットT₁～T₄における動作に分けて示している。

【0027】

まず、第1の各タイムスロットT₁では、図17に示した第1～第4のインプットモジュール151₁～151₄にフレームパルス152が入力されると、第1のインプットモジュール151₁が次フレームのタイムスロットT₁における第1の入力ポートの接続要求情報を最初に決定する。最初の決定であるので、出力ポート予約済み情報162₁は、第1の出力ポートから第4の出力ポートまで順番に（0，0，0，0）となっている。第1の入力ポートの接続要求156₁が第1の出力ポートから第4の出力ポートまで順番に（0，1，0，1）であるとする。第1のインプットモジュール151₁が第2および第4の出力ポートの中から第2の出力ポートを選択したとする。この場合第1のインプットモジュール151₁は次フレームのタイムスロットT₁の接続許可情報157₁として第2の出力ポートを記憶し、出力ポート予約済み情報161₁を（0，1，0，0）として第2のインプットモジュール151₂に通知する。

【0028】

続けてタイムスロット T_2 で第2のインプットモジュール 151_2 が次フレームのタイムスロット T_2 に対する第2の入力ポートの接続許可情報 157_2 を決定する。第2のインプットモジュール 151_2 は出力ポート予約済情報 162_2 (0, 1, 0, 0)を受信する。第2の入力ポートの接続要求 156_2 が(0, 1, 1, 1)であるとする、第2の出力ポートに対する割当は行えない。このため、第2のインプットモジュール 151_2 は第2の出力ポートに対する接続要求は拒否し、第3または第4の出力ポートの接続要求から許可する出力ポートを選択する。ここで第3の出力ポートを選択したとすると、第2のインプットモジュール 151_2 は接続許可情報 157_2 として第3の出力ポートを記憶し、出力ポート予約済情報 161_2 を(0, 1, 1, 0)として第3のインプットモジュール 151_3 に通知する。

【0029】

以下、タイムスロット T_3 で第3のインプットモジュール 151_3 が、またタイムスロット T_4 で第4のインプットモジュール 151_4 が次フレームのタイムスロット T_1 の接続許可情報 157_3 、 157_4 を決定する。タイムスロット T_4 が終了した時点で各モジュールは次フレームのタイムスロット T_1 における接続許可情報 157_4 を有しているので、次フレームのタイムスロット T_1 における 4×4 の接続許可情報が確定することとなる。

【0030】

以上説明した処理手順では、タイムスロット T_1 で第1のインプットモジュール 151_1 以外のモジュールがそれぞれ異なるタイムスロットに対する「予約」を開始するようになっている。例えば、第2のインプットモジュール 151_2 は次フレームのタイムスロット T_4 に、第3のインプットモジュール 151_3 は次フレームのタイムスロット T_3 に、また第4のインプットモジュール 151_4 は次フレームのタイムスロット T_2 に対する「予約」を開始する。

【0031】

各モジュールは当該タイムスロットにて各予約タイムスロットの処理を行い、該当する予約タイムスロットの出力ポート予約情報 161 をそれぞれ次段のインプットモジュールに転送することにより、各インプットモジュール $151_1 \sim 1$

5 1₄とも未動作の時間がないようにスケジューリング処理を実行することができる。タイムスロットT₄が終了した時点で、各モジュールは次フレームのタイムスロットT₁からT₄までの接続許可情報1 5 7を有している。従って、次フレームのタイムスロットT₁からT₄までの4×4の接続許可情報1 5 7₁～1 5 7₄が確定することとなる。

【0 0 3 2】

図1 9は各モジュールの予約（接続許可決定）順序を示したものである。この図は上の例で説明した4×4スケジューラの場合を示している。横軸は時間であり4タイムスロットで1つのフレームを構成している。縦軸にはモジュールの物理番号1 5 5₁～1 5 5₄が並べられている。出力ポート予約済情報の転送方向は、インプットモジュールの物理番号で示すと、物理番号1 5 5₁、物理番号1 5 5₂、物理番号1 5 5₃、物理番号1 5 5₄、物理番号1 5 5₁、物理番号1 5 5₂、……の順序となる。マトリックス内の数字は予約をする次フレーム内のタイムスロットの番号を指す。

【0 0 3 3】

図1 9に示されているように、物理番号1 5 5₁のインプットモジュールは、フレームの先頭（タイムスロットT₁）で次フレームのタイムスロットT₁の予約から開始する。以下同様に、物理番号1 5 5₂のインプットモジュールはタイムスロットT₄の予約から開始し、物理番号1 5 5₃のインプットモジュールはタイムスロットT₃の予約から開始する。物理番号1 5 5₄のインプットモジュールはタイムスロットT₂の予約から開始する。

【0 0 3 4】

【発明が解決しようとする課題】

【0 0 3 5】

以上説明したパケットスイッチでは、入力ポート側をパイプライン化することで、どの出力ラインに出力するか調停をスイッチスケジューラ1 2 3（図1 6参照）で調停している。この結果、たとえば図1 7に示す第1のインプットモジュール1 5 1₁が、図1 9に示したように第1のタイムスロットT₁に出力ポートの予約を開始し、以後、第2のインプットモジュール1 5 1₂以降のインプット

モジュールに対して順次各タイムスロットごとに出力ポート予約済み情報 1 6 1 を受け渡していくとき、第 2 のインプットモジュール 1 5 1₂ は第 2 のタイムスロット T₂ に出力ポートの予約を開始し、以後、第 3 のインプットモジュール 1 5 1₃ 以降のインプットモジュールに対して順次各タイムスロットごとに出力ポート予約済み情報 1 6 1 を受け渡していくことになる。第 3 のインプットモジュール 1 5 1₃ および第 4 のインプットモジュール 1 5 1₄ も、それぞれ更に 1 タイムスロットあるいは 2 タイムスロットずらした点から出力ポートの予約を開始し、以降のインプットモジュールに対して順次各タイムスロットごとに出力ポート予約済み情報 1 6 1 を受け渡していく。このような並行的な処理によって、パイプライン化しなかった場合と比べると、入力ポートおよび出力ポートの増大に対して処理の高速化が可能となる。

【 0 0 3 6 】

以上、特願平 1 1 - 3 5 5 3 8 2 号記載の技術について詳細に説明したが、同様の試みはたとえば特願平 1 1 - 2 3 4 1 5 8 号でも行われている。

【 0 0 3 7 】

しかしながら、このような入力ポート側のパイプライン化処理だけでは入力ポートと出力ポートの増加に伴い出力ポート側について調停を行わなければならないデータ処理量が出力ポートの数の増加に比例して増加する。転送するデータの大容量化が顕著となっている現在、データ伝送の高速化が大きな社会的要請となっており、パケットスイッチの転送単位としての固定長セルのスイッチングに遅延が生じることは避けなければならず、むしろ処理速度の高速化が急務となっている。

【 0 0 3 8 】

そこで本発明の目的は、入力ポート側のパイプライン化処理を実現した場合よりも更に高速処理が可能なパケット交換装置を提供することにある。

【 0 0 3 9 】

【課題を解決するための手段】

請求項 1 記載の発明では、(イ) 複数の入力ポートのそれぞれに用意され、入力されるそれぞれのパケットを宛先の出力ポートに対応した格納場所に順に格納

するバッファメモリと、（ロ）それぞれの入力ポートのバッファメモリに格納されたパケットについての転送要求の有無を表わした情報を、予め全入力ポートと全出力ポートを重複しないように等しい数ずつ選択してなる入力ポートと出力ポートの組み合わせからなる $M \times M$ のマトリックス状に配置された複数のグループのうちの該当する入力ポートと出力ポートに対応した1つずつのグループに割り振るグループ別割振り手段と、（ハ）このグループ別割振り手段によって割り振られたそれぞれのグループの中で、 $M \times M$ のマトリックス上で行および列方向に競合が生じない M 個のグループを選択したときのこれらのグループを $M \times M$ のマトリックスの行方向と列方向の2方向にパイプライン処理して固定長セルの転送要求を表わした情報の予約をグループ単位で行う予約処理手段と、（ニ）この予約処理手段の予約処理結果に対応させてバッファメモリに格納されたパケットの入出力を設定して転送要求のあったパケットの転送を行うスイッチとをパケット交換装置に具備させる。

【 0 0 4 0 】

すなわち請求項1記載の発明では、パケットの送られてくる入力ポートとこれらのパケットの出力先としての出力ポートを M 個ずつに分けてなる $M \times M$ のマトリックス状に配置されたグループにそれぞれのパケットの入出力の関係の情報としての転送要求の有無を表わした情報を割り振っておく。これらのグループは、予め全入力ポートと全出力ポートを重複しないように等しい数ずつ選択してなる入力ポートと出力ポートの組み合わせからなっているので、マトリックスの行方向にも列方向にもパイプライン処理が可能である。そこで、予約処理手段は両方向にパイプライン処理を行って、転送要求の予約と結果の受け渡しをグループごとに受け継いでいく。スイッチはこのようにして得られた予約処理結果に対応させてバッファメモリに格納されたパケットの入出力を設定して転送要求のあったパケットをそれぞれの宛先としての出力ポートに転送させる。このように請求項1記載の発明ではマトリックスの行方向と列方向、すなわち入力ポート側と出力ポート側の双方にパイプライン処理を行うことで、パケットの転送処理の高速化を可能にしている。したがって、ポートの数の増大に十分対応することができる。

【 0 0 4 1 】

請求項 2 記載の発明では、（イ）複数の入力ポートに入力されるそれぞれのパケットを所定の固定長セルに分解する固定長セル分解手段と、（ロ）それぞれの入力ポートに対応して配置され、この固定長セル分解手段によって分解されたそれぞれのセルを宛先の出力ポートに対応した格納場所に順に格納する論理キューと、（ハ）それぞれの入力ポートの論理キューに格納された固定長セルについての転送要求の有無を表わした情報を、予め全入力ポートと全出力ポートを重複しないように等しい数ずつ選択してなる入力ポートと出力ポートの組み合わせからなる $M \times M$ のマトリックス状に配置された複数のグループのうちの該当する入力ポートと出力ポートに対応した 1 つずつのグループに割り振るグループ別割り振り手段と、（ニ）このグループ別割り振り手段によって割り振られたそれぞれのグループの中で、 $M \times M$ のマトリックス上で行および列方向に競合が生じない M 個のグループを選択したときのこれらのグループを $M \times M$ のマトリックスの行方向と列方向の 2 方向にパイプライン処理して固定長セルの転送要求を表わした情報の予約をグループ単位で行う予約処理手段と、（ホ）この予約処理手段の予約処理結果に対応させて論理キューに格納された固定長セルの入出力を設定して転送要求のあった固定長セルの転送を行うスイッチと、（ヘ）このスイッチによってそれぞれの出力側に得られた固定長セルをパケットに組み立てるパケット組立手段とをパケット交換装置に具備させる。

【 0 0 4 2 】

すなわち請求項 2 記載の発明では、スイッチの処理単位としての所定の固定長セルにパケットを分解する一方、パケットの送られてくる入力ポートとこれらのパケットの分解後の固定長セルの出力先としての出力ポートを M 個ずつに分けてなる $M \times M$ のマトリックス状に配置されたグループにそれぞれの固定長セルの入出力の関係の情報としての転送要求の有無を表わした情報を割り振っておく。これらのグループは、予め全入力ポートと全出力ポートを重複しないように等しい数ずつ選択してなる入力ポートと出力ポートの組み合わせからなっているので、マトリックスの行方向にも列方向にもパイプライン処理が可能である。そこで、予約処理手段は両方向にパイプライン処理を行って、固定長セルごとの転送要求

の予約と結果の受け渡しをグループごとに受け継いでいく。スイッチはこのようにして得られた予約処理結果に対応させてバッファメモリに格納された固定長セルの入出力を設定して転送要求のあった固定長セルをそれぞれの宛先としての出力ポートに転送させる。このように請求項 2 記載の発明ではマトリックスの行方向と列方向、すなわち入力ポート側と出力ポート側の双方にパイプライン処理を行うことで、固定長セルの転送処理の高速化を可能にしている。したがって、ポートの数の増大に十分対応することができる。しかも、パケットを固定長セルに分解して処理するので、可変長のパケットについても高速処理が可能になる。

【 0 0 4 3 】

請求項 3 記載の発明では、請求項 1 または請求項 2 記載のパケット交換装置で、予約処理手段は、1つのグループの予約から次のグループへの予約結果の転送までを1つのタイムスロットで行うことを特徴としている。

【 0 0 4 4 】

すなわち請求項 3 記載の発明では、予約処理手段が、1つのグループの予約から次のグループへの予約結果の転送までを1つのタイムスロットで処理することになっているので、マトリックスの行方向にも列方向にも処理を行う本発明で全体としての処理が規則正しく進行することになる。

【 0 0 4 5 】

請求項 4 記載の発明では、請求項 1 または請求項 2 記載のパケット交換装置で、予め全入力ポートと全出力ポートを重複しないように等しい数ずつ選択してなる入力ポートと出力ポートの組み合わせは、ダイアゴナル・サービス・パターンであり、そのM個の組み合わせからなるダイアゴナル・モジュール群を用いてパイプライン処理が行われることを特徴としている。

【 0 0 4 6 】

すなわち請求項 4 記載の発明では、 $M \times M$ のグループに対してM個からなるダイアゴナル・サービス・パターンをM通り用意することで、いずれのタイムスロットでも1組のダイアゴナル・サービス・パターンによる予約処理が開始するので、効率的な予約処理が可能になる。また、 $M \times M$ 個の全グループがダイアゴナル・サービス・パターンの構成となるので、予約処理の優先度が結果的に均等に割り振

られることになり、入力ポートによって優劣が生じるといった問題を生じさせない。

【0047】

請求項5記載の発明では、請求項4記載のパケット交換装置で、予約処理手段は、予め定めた未来の所定のタイムスロットに対して、ダイアゴナル・サービス・パターンに該当するグループから該当する出力ポートに対して順次転送要求予約を行うことを特徴としている。

【0048】

すなわち請求項5記載の発明では、予め定めた未来の所定のタイムスロットに対して、ダイアゴナル・サービス・パターンに該当するグループを起点として予約を行うことにしているので、これら未来のタイムスロットを重複しないように割り当てることで、効率的な処理を実現することができる。また、これらのグループを必要な個数連結して順次予約処理を実行させることで、スケラブルなスケジューラを実現することができる。

【0049】

【発明の実施の形態】

【0050】

【実施例】

以下実施例につき本発明を詳細に説明する。

【0051】

図1は本発明の一実施例におけるパケット交換装置の入力バッファ型スイッチのスイッチ部およびその周辺の構成を表わしたものである。入力バッファ型スイッチ全体の構成は、先に示した図15と同様である。本実施例のスイッチ部201は、スイッチ素子としての $N \times N$ のクロスポイントスイッチ202と、スイッチスケジューラ203とによって構成されている。スイッチ部201には N 本の入力ライン204₁、204₂、……204_Nと接続された入力ポートと、同じく N 本の出力ライン205₁、205₂、……205_Nと接続された出力ポートが配置されている。各入力ポートは図15で示したVOQ106と同様のVOQ206₁、206₂、……206_Nを備えている。

【 0 0 5 2 】

転送対象のパケットは、図 1 5 に示したパケット生成分解部 (PAD) 1 0 5 と同一の回路部分でクロスポイントスイッチ 2 0 2 の転送単位の固定長セルに分解されている。この固定長セル単位での転送時間を 1 タイムスロットと呼んでいる。N 本の入力ライン $2 0 4_1$ 、 $2 0 4_2$ 、…… $2 0 4_N$ は、M 個 (ここで数値 M は 2 以上の整数であり、かつ整数 N の約数。) のグループ 2 1 1 に均等に分割されている。N 本の出力ライン $2 0 5_1$ 、 $2 0 5_2$ 、…… $2 0 5_N$ も同様に M 個のグループ 2 1 2 に均等に分割されている。これら M 個の入力グループ 2 1 1 と M 個の出力グループ 2 1 2 の組み合わせは、 M^2 となる。スイッチスケジューラ 2 0 3 は、これら M^2 の組み合わせに対してスケジューリング・モジュールを用意している。

【 0 0 5 3 】

図 2 は、これら M^2 個のスケジューリング・モジュールとその働きを示したものである。スケジューリング・モジュール $2 2 1_1$ 、…… $2 2 1_{MM}$ は、この図 2 で横軸方向すなわち図 1 に示す出力ライン $2 0 5_1$ 、 $2 0 5_2$ 、…… $2 0 5_N$ の方向に M 個ずつの組 ($2 2 1_1$ 、 $2 2 1_2$ 、…… $2 2 1_M$)、($2 2 1_{M+1}$ 、 $2 2 1_{M+2}$ 、…… $2 2 1_{2M}$)、……を、縦軸方向すなわち図 1 で入力ライン $2 0 4_1$ 、 $2 0 4_2$ 、…… $2 0 4_N$ の方向に M 回繰り返して配置したものと等しい。この場合、最後の組は図 2 で一番下に示した組 ($2 2 1_{M(M-1)+1}$ 、 $2 2 1_{M(M-1)+2}$ 、…… $2 2 1_{MM}$) となる。

【 0 0 5 4 】

それぞれのスケジューリング・モジュール $2 2 1_1$ 、…… $2 2 1_{MM}$ は、VOQ $2 0 6_1$ 、 $2 0 6_2$ 、…… $2 0 6_N$ のうちの対応するものからのパケット (固定長セル) 転送要求に基づいて、予め定められた未来のタイムスロットにおける $(N/M) \times (N/M)$ 通りのスケジューリングを実行する。そして、それぞれのスケジューリングの結果として予約済み出力ポート情報 $2 3 1_1$ 、 $2 3 1_2$ 、…… $2 3 1_M$ と、予約済み入力ポート情報 $2 3 2_1$ 、 $2 3 2_2$ 、…… $2 3 2_M$ とを、それぞれ予め定められた次段のスケジューリング・モジュールに転送するようになっている。

【0055】

この図2に示した例では、パイプライン処理で情報が一巡するためのこのような転送処理を、図で横方向に隣接するそれぞれのスケジューリング・モジュール221の間で一巡させる処理と、図で縦方向に隣接するそれぞれのスケジューリング・モジュール221の間で一巡させる処理について示している。このような処理におけるスケジューリング・モジュール221同士の接続順序は、これら隣接するもの同士に限る必要はない。すなわち、横方向あるいは縦方向のそれぞれについて、情報が一巡するような任意の順序で各スケジューリング・モジュール221の接続順序を定めることができる。この接続順序は、後に説明するダイアゴナル・モジュール群の選び方によって規定されることになる。スケジューリングおよび情報転送の処理は、それぞれ1タイムスロット以内に実行される。次段のスケジューリング・モジュールでは、対応する論理キューからのパケット転送要求と、前段からの予約済み入出力ポート情報に基づいて、衝突を回避するような予約が行われる。

【0056】

図3は、VOQの1つを拡大して図解したものである。VOQ206₁、206₂、……206_Nはそれぞれ同一構成となっているので、ここでは第1のVOQ206₁を例示的に示すものとする。第1のVOQ206₁は、図1に示した第1の入力ライン204₁に対応しており、図15で示したパケット生成分解部(PAD)105を経た固定長セル241をN本の出力ライン205₁、205₂、……205_Nに対応させた論理キュー(バッファメモリ)242₁、242₂、……242_Nに格納するようになっている。この図で各論理キュー242₁、242₂、……242_Nによって固定長セル241の格納された部分の長さが異なって表示されているのは、格納された固定長セル241の数がそれぞれ異なっているためである。

【0057】

図3に示した第1のVOQ206₁は、全部でMのグループに分けられている。図2では全部でM²個のスケジューリング・モジュール221₁、……221_{MM}を表わしたが、これらは出力ライン205₁、205₂、……205_Nの方向にM

グループに分かれている。図 3 ではこれと同様に第 1 の $VOQ\ 206_1$ の各論理キュー 242_1 、 242_2 、…… 242_N を M グループに分けている。図 3 で太線で示した第 1 のグループ 244_1 は、第 1 ～ 第 N/M の論理キュー 242_1 、 242_2 、…… $242_{N/M}$ に対応している。第 2 のグループ 244_2 は、第 $N/M+1$ ～ 第 $2N/M$ の論理キュー $242_{N/M+1}$ 、 $242_{N/M+2}$ 、…… $242_{2N/M}$ に対応している。以下同様にして、第 M のグループ 244_M は、第 $N-N/M+1$ ～ 第 N の論理キュー $242_{N-N/M+1}$ 、 $242_{N-N/M+2}$ 、…… 242_N に対応している。

【 0 0 5 8 】

以上のような構成のパケットスイッチの動作を次に説明する。ただし、説明を分かりやすくするために、以上説明した数値 N が “16” で、数値 M が “4” という比較的小さな値の場合の例を挙げる。これは、 16×16 のクロスポイントスイッチの縦方向と横方向をそれぞれ 4 つのグループに区分けした場合と等しい。

【 0 0 5 9 】

図 4 はこの 16×16 のクロスポイントスイッチにおけるスケジューリング・モジュールを表わしたものである。この例では、16 本の入力を 4 本ずつ 4 グループに分割することになっている。16 本の出力も同様に 4 本ずつ 4 グループに分割される。入力が 4 グループで出力が 4 グループの場合には、これらを組み合わせたとき 16 対のグループが存在する。これらを第 1 ～ 第 16 のスケジューリング・モジュール $221_1 \sim 221_{16}$ で表わすものとする。

【 0 0 6 0 】

第 1 ～ 第 16 のスケジューリング・モジュール $221_1 \sim 221_{16}$ には、それぞれの処理対象とする入力ポートから出力ポートへの転送要求が入力される。この例では合計 16 本の入力ライン 204_1 、 204_2 、…… 204_{16} と、同じく 16 本の出力ライン 205_1 、 205_2 、…… 205_{16} が存在しているので、これらがそれぞれ 4 つのグループに分けられて第 1 ～ 第 16 のスケジューリング・モジュール $221_1 \sim 221_{16}$ に入力される。たとえば、第 2 の入力ポートから入力された固定長セルの宛先が第 3 の出力ポートであるとする。第 2 の入力ポートは入力ポートを 4 つのグループに分けたときの図 4 における縦方向に示した第 1 の

グループに属し、第3の出力ポートは図4における横方向に示した出力ポートを4つのグループに分けたときの第1のグループに属する。したがってこの転送要求は図4における、縦方向も横方向も最初のグループとして位置付けられている第1のスケジューリング・モジュール221₁に入力されることになる。

【0061】

また、第7の入力ポートから第11の出力ポートへの転送要求については、第7の入力ポートが入力ポートを4つのグループに分けたときの図4における縦方向に示した第2のグループに属し、第11の出力ポートは図4における横方向に示した出力ポートを4つのグループに分けたときの第3のグループに属する。したがってこの転送要求は図4における、縦方向が2つ目のグループで横方向が3つ目のグループとして位置付けられている第7のスケジューリング・モジュール221₇に入力されることになる。

【0062】

図5は、第1のスケジューリング・モジュールのマトリックスの内部の信号配置を一例として示したものである。この図に例示したように、図4の各スケジューリング・モジュール221₁～221₁₆の中に示したマトリックスには、それぞれの論理キューから受信したパケット転送要求の有無を示す信号“0”または信号“1”が格納される。ここで、信号“0”は論理キューからのパケット転送の要求がないことを示しており、信号“1”は論理キューからのパケット転送の要求があることを示している。

【0063】

本実施例では、ダイアゴナル・サービス・パターンを構成するダイアゴナル・モジュール群として、図4で太枠で示した第1のスケジューリング・モジュール221₁、第6のスケジューリング・モジュール221₆、第11のスケジューリング・モジュール221₁₁および第16のスケジューリング・モジュール221₁₆の4つを選択している。これらはマトリックス上で各入力ポートおよび出力ポートの競合が生じないような組み合わせとなっている。すなわち、第1のスケジューリング・モジュール221₁では第1～第4の入力ポートからの転送要求の有無を示す信号が格納されるようになっており、これは第5～第8の入力ポートを

扱う第 6 のスケジューリング・モジュール 221_6 や、第 9 ～第 12 の入力ポートを扱う第 11 のスケジューリング・モジュール 221_{11} や、第 13 ～第 16 の入力ポートを扱う第 16 のスケジューリング・モジュール 221_{16} と入力ポート間で競合しない。同様に、第 1 のスケジューリング・モジュール 221_1 では第 1 ～第 4 の出力ポートへの転送要求の有無を示す信号が格納されるようになっており、これは第 5 ～第 8 の出力ポートを扱う第 6 のスケジューリング・モジュール 221_6 や、第 9 ～第 12 の出力ポートを扱う第 11 のスケジューリング・モジュール 221_{11} や、第 13 ～第 16 の出力ポートを扱う第 16 のスケジューリング・モジュール 221_{16} と出力ポート間でも競合しない。

【 0 0 6 4 】

このようなダイアゴナル・サービス・パタンを構成するダイアゴナル・モジュール群は、時間的に重複して他にも 3 組存在することができる。これらは (S_2 、 S_7 、 S_{12} 、 S_{13})、(S_3 、 S_8 、 S_9 、 S_{14})、(S_4 、 S_5 、 S_{10} 、 S_{15}) の各組である。但しここでは表記を単純化するために、スケジューリング・モジュール 221_x を単に S_x と表記し、ダイアゴナル・モジュール群ごとに括弧で括って示している。

【 0 0 6 5 】

図 4 に示した各スケジューリング・モジュール $221_1 \sim 221_{16}$ は、1 つの固定長セルの転送時間内に、予め定められたある未来のタイムスロットについての 4×4 のスケジューリングを実行する。そして、その結果得られる予約済み出力ポート情報 $261 \sim 264$ を図の縦方向（縦の矢印方向）の次段に位置するスケジューリング・モジュール 221 に転送する。また、同時に予約済み入力ポート情報 $265 \sim 268$ を図の横方向（横の矢印方向）の次段に位置するスケジューリング・モジュール 221 にも転送する。たとえば、第 1 のスケジューリング・モジュール 221_1 について説明すると、予約済み出力ポート情報 261 を第 13 のスケジューリング・モジュール 221_{13} に転送すると共に、予約済み入力ポート情報 265 を第 2 のスケジューリング・モジュール 221_2 に転送する。なお、本実施例では図 4 で太枠で示した第 1、第 6、第 11 および第 16 のスケジューリング・モジュール 221_1 、 221_6 、 221_{11} 、 221_{16} は、対応する

論理キューから転送要求の有無を示す信号を格納されるので、予約はこれらのスケジューリング・モジュール 221_1 、 221_6 、 221_{11} 、 221_{16} から開始されることになる。

【0066】

各スケジューリング・モジュール $221_1 \sim 221_{16}$ は、対応する論理キューからの転送要求と前段からの予約済み入出力ポート情報に基づいて、次のタイムスロットについて 4×4 のスケジューリングを実行し、その結果を更に（縦横両方向の）次段のスケジューリング・モジュール 221 に転送することになる。このようにして、4タイムスロットかけて、このような処理が第1～第16のスケジューリング・モジュール $221_1 \sim 221_{16}$ のすべてで行われることになる。これにより、ある未来のタイムスロットについてのスケジューリングが完成する。

【0067】

以上の処理が実行されていくためには、幾つかの条件が必要である。まず、1つのスケジューリング・モジュール 221 内における 4×4 のスケジューリングについての処理は、次段への情報転送を含めて1タイムスロット内で完了しなければならない。また、4入力の論理キューの予約機会が均等になることである。

【0068】

以上の条件を満たすものであれば、各種のアルゴリズムを本実施例のケット交換装置に適用することができる。たとえば入力×出力の転送要求をマトリックス状に並べて検索を行い、割り当てを行うアルゴリズムが特開平6-233337号公報あるいは特開平9-326828号公報に開示されている。これらの技術も前記した条件を満たすものであれば、本実施例のケット交換装置に使用することができる。

【0069】

図6～図8は、各スケジューリング・モジュールのスケジューリングの様子を時間の変化と共に表わしたものである。図6は最初のタイムスロット T_1 からの処理を表わしており、図7、図8は順次時間が経過した状態を表わしている。これらの図を用いて、本実施例の第1～第16のスケジューリング・モジュール2

2 1₁ ~ 2 2 1₁₆ (図では S₁ ~ S₁₆ と略記する。) によるスケジューリングを具体的に説明する。

【 0 0 7 0 】

すでに説明したように本実施例では、ある未来のタイムスロットの割り当てとしての入力と出力の組み合わせの決定を、図 2 に示した予約済み出力ポート情報 2 3 1 あるいは予約済み入力ポート情報 2 3 2 の流れる順路で予約を行うようにしている。たとえば図 4 に太枠で記した第 1、第 6、第 1 1 および第 1 6 のスケジューリング・モジュール 2 2 1₁、2 2 1₆、2 2 1₁₁、2 2 1₁₆ (S₁、S₆、S₁₁、S₁₆) は、本実施例でダイアゴナル・モジュール群を構成しているため、最初のタイムスロット T₁ で、これから 4 つ後のタイムスロット T₅ における入力と出力の組み合わせの予約を 1 番目に行う権利を有する。

【 0 0 7 1 】

たとえば、タイムスロット T₁ で第 1 ~ 第 4 の入力ポートから第 1 ~ 第 4 の出力ポートへ固定長セルの転送要求があったとする。これらの入出力ポートに対応する第 1 の分散スケジューリング・モジュール 1 2 7₁ (S₁) は、これらの要求が競合するような場合には、その時点で各要求を調停する。一例としては、第 1 ~ 第 4 の入力ポートがすべて第 1 の出力ポートへ固定長セルを転送する要求を行っていたとすると、このうちの 1 つの入力ポートしか第 1 の出力ポートへ固定長セルを転送することはできない。したがって、これら競合した入力ポートの要求をいずれか 1 つに選択するための調停が行われる。

【 0 0 7 2 】

同様に第 5 ~ 第 8 の入力ポートから第 5 ~ 第 8 の出力ポートへ固定長セルの転送要求があった場合には、第 6 の分散スケジューリング・モジュール 1 2 7₆ (S₆) はそれらの要求を調停して出力を予約する。また、第 9 ~ 第 1 2 の入力ポートから第 9 ~ 第 1 2 の出力ポートへ固定長セルの転送要求があった場合には、第 1 1 の分散スケジューリング・モジュール 1 2 7₁₁ (S₁₁) がそれらの要求を調停して出力を予約する。更に、第 1 3 ~ 第 1 6 の入力ポートから第 1 3 ~ 第 1 6 の出力ポートへ固定長セルの転送要求があった場合には、第 1 6 の分散スケジューリング・モジュール 1 2 7₁₆ (S₁₆) がそれらの要求を調停して出力を予約

することになる。これら、ダイアゴナル・モジュール群を構成する第1、第6、第11および第16のスケジューリング・モジュール 221_1 、 221_6 、 221_{11} 、 221_{16} (S_1 、 S_6 、 S_{11} 、 S_{16}) の処理は、図6における1つの処理群281内での最初のタイムスロット T_1 の処理として表わしている。

【0073】

このようにして、ダイアゴナル・モジュール群を構成する第1、第6、第11および第16のスケジューリング・モジュール 221_1 、 221_6 、 221_{11} 、 221_{16} (S_1 、 S_6 、 S_{11} 、 S_{16}) が最初に出力ポートの予約を行うと、次のタイムスロット T_2 で次のスケジューリング・モジュールに予約処理が引き継がれる。このために、タイムスロット T_1 で得られた予約済み出力ポート情報261～264は、それぞれ図4に示す縦方向（縦の矢印方向）における次段としての第13、第2、第7および第12のスケジューリング・モジュール 221_{13} 、 221_2 、 221_7 、 221_{12} (S_{13} 、 S_2 、 S_7 、 S_{12}) に転送される。また、同時に予約済み入力ポート情報265～268は横方向（横の矢印方向）における次段としての第2、第7、第12および第13のスケジューリング・モジュール 221_2 、 221_7 、 221_{12} 、 221_{13} (S_2 、 S_7 、 S_{12} 、 S_{13}) にも転送されることになる。

【0074】

この結果、タイムスロット T_2 では、第2、第7、第12および第13のスケジューリング・モジュール 221_2 、 221_7 、 221_{12} 、 221_{13} (S_2 、 S_7 、 S_{12} 、 S_{13}) でタイムスロット T_5 の予約が行われることになる。この際には、第6、第11、第16および第1のスケジューリング・モジュール 221_6 、 221_{11} 、 221_{16} 、 221_1 (S_6 、 S_{11} 、 S_{16} 、 S_1) から縦方向（縦の矢印方向）に転送されてきた予約済み出力ポート情報262～264、261と、第1、第6、第11および第16のスケジューリング・モジュール 221_1 、 221_6 、 221_{11} 、 221_{16} (S_1 、 S_6 、 S_{11} 、 S_{16}) から横方向（横の矢印方向）に転送されてきた予約済み入力ポート情報265～268を参照して、グループ内の論理キューからの要求に基づいて予約を行う。当然ながら、先のタイムスロット T_1 で予約を行った入出力関係について予約を行うことはできず、未予約とな

っている入出力の組み合わせの中から新たな予約を行うことになる。

【0075】

このようにして第2、第7、第12および第13のスケジューリング・モジュール221₂、221₇、221₁₂、221₁₃(S₂、S₇、S₁₂、S₁₃)でタイムスロットT₅の予約が行われると、タイムスロットT₃で次のスケジューリング・モジュールに予約処理が引き継がれる。このために、タイムスロットT₂で得られた予約済み出力ポート情報262~264、265は、それぞれ図4に示す縦方向(縦の矢印方向)における次段としての第14、第3、第8および第9のスケジューリング・モジュール221₁₄、221₃、221₈、221₉(S₁₄、S₃、S₈、S₉)に転送される。また、同時に予約済み入力ポート情報265~268は横方向(横の矢印方向)における次段としての第3、第8、第9および第14のスケジューリング・モジュール221₃、221₈、221₉、221₁₄(S₃、S₈、S₉、S₁₄)に転送される。そして、第3、第8、第9および第14のスケジューリング・モジュール221₃、221₈、221₉、221₁₄(S₃、S₈、S₉、S₁₄)でタイムスロットT₅の予約が行われる。

【0076】

以下同様にして、タイムスロットT₃の次のタイムスロットT₄では、第4、第5、第10および第15のスケジューリング・モジュール221₄、221₅、221₁₀、221₁₅(S₄、S₅、S₁₀、S₁₅)でタイムスロットT₅の予約が行われる。このようにして、図7の1つの処理群281における処理が終了すると、第1~第16のスケジューリング・モジュール221₁~221₁₆(S₁~S₁₆)のすべてについて、すなわち第1~第4の入力ポートから第1~第4の出力ポートに出力される固定長セルの全転送要求についてのタイムスロットT₅の予約が完了することになる。

【0077】

ところで、図4で太枠で示したダイアゴナル・サービス・ボタンを構成する第1、第6、第11および第16のスケジューリング・モジュール221₁、221₆、221₁₁、221₁₆(S₁、S₆、S₁₁、S₁₆)の4つに着目してみると、これらはタイムスロットT₅の予約のための最初のタイムスロットT₁における処

理で、それぞれの分担する予約処理を完了させてしまう。したがって、これ以後のタイムスロット T_4 までの時間を空き時間として何らの処理も行わないのは効率的ではない。また、これらの処理結果を引き継いでタイムスロット T_2 で予約処理を行う第2、第7、第12および第13のスケジューリング・モジュール221₂、221₇、221₁₂、221₁₃ (S_2 、 S_7 、 S_{12} 、 S_{13}) について見てみると、最初のタイムスロット T_1 で何の処理も行わずに待機しているのは時間の無駄であるし、同様にタイムスロット T_2 で予約処理を行った後に次に予約処理を行うまで新たに3タイムスロット分だけ何らの処理も行わないのも時間の無駄である。第3、第8、第9および第14のスケジューリング・モジュール221₃、221₈、221₉、221₁₄ (S_3 、 S_8 、 S_9 、 S_{14}) および第4、第5、第10および第15のスケジューリング・モジュール221₄、221₅、221₁₀、221₁₅ (S_4 、 S_5 、 S_{10} 、 S_{15}) についても、同様に時間に対する稼働率が4分の1であるという問題が生じる。

【0078】

そこで、本実施例のスケジューリング処理では、図6に示した処理群281の他に、その上の部分に示した処理群282～284が並列して動作するようになっている。すなわち、処理群282について見てみると、最初のタイムスロット T_1 で第2、第7、第12および第13のスケジューリング・モジュール221₂、221₇、221₁₂、221₁₃ (S_2 、 S_7 、 S_{12} 、 S_{13}) がタイムスロット T_6 における予約処理を行い、その結果を第3、第8、第9および第14のスケジューリング・モジュール221₃、221₈、221₉、221₁₄ (S_3 、 S_8 、 S_9 、 S_{14}) に渡してタイムスロット T_2 でタイムスロット T_6 における予約処理を行わせる。以下同様にしてタイムスロット T_3 では第4、第5、第10および第15のスケジューリング・モジュール221₄、221₅、221₁₀、221₁₅ (S_4 、 S_5 、 S_{10} 、 S_{15}) でタイムスロット T_6 における予約処理が行われ、タイムスロット T_4 では第1、第6、第11および第16のスケジューリング・モジュール221₁、221₆、221₁₁、221₁₆ (S_1 、 S_6 、 S_{11} 、 S_{16}) でタイムスロット T_6 における予約処理が行われることになる。

【0079】

処理群 283 については、最初のタイムスロット T_1 で第 3、第 8、第 9 および第 14 のスケジューリング・モジュール 221_3 、 221_8 、 221_9 、 221_{14} (S_3 、 S_8 、 S_9 、 S_{14}) がタイムスロット T_7 における予約処理を行い、以下同様にして以降のタイムスロットでそれぞれ対応するスケジューリング・モジュールがタイムスロット T_7 における予約処理を行う。最後の処理群 284 では、同様にしてそれぞれ同一のタイムスロットで処理の行われなかったスケジューリング・モジュールがタイムスロット T_8 における予約処理を行うことになる。

【0080】

図 7 および図 8 に示した各処理群 285 ~ 292 についても同様である。このように 4×4 のスケジューリングは、図 4 で太枠で示したダイアゴナル・サービス・パターンを構成する第 1、第 6、第 11 および第 16 のスケジューリング・モジュール 221_1 、 221_6 、 221_{11} 、 221_{16} (S_1 、 S_6 、 S_{11} 、 S_{16}) の 4 つのみが予約処理に優先するのではなく、それぞれ起点となるタイムスロットを異にした第 2、第 7、第 12 および第 13 のスケジューリング・モジュール 221_2 、 221_7 、 221_{12} 、 221_{13} (S_2 、 S_7 、 S_{12} 、 S_{13})、第 3、第 8、第 9 および第 14 のスケジューリング・モジュール 221_3 、 221_8 、 221_9 、 221_{14} (S_3 、 S_8 、 S_9 、 S_{14})、第 4、第 5、第 10 および第 15 のスケジューリング・モジュール 221_4 、 221_5 、 221_{10} 、 221_{15} (S_4 、 S_5 、 S_{10} 、 S_{15}) の各組も優先の機会を与えられる。すなわち、第 1 ~ 第 16 のスケジューリング・モジュール $221_1 \sim 221_{16}$ が、4 つのダイアゴナル・サービス・パターンによって平等の処理を受けることになる。

【0081】

図 9 ~ 図 14 はスケジューリング・モジュールが 4×4 の合計 16 個で構成される場合のダイアゴナルモジュール群の選び方とモジュール間の接続の順序を示したものである。各 4×4 マトリックス内の黒丸がダイアゴナル・サービス・パターンを構成している。また、矢印で連結した横 1 行の 4 つのダイアゴナル・サービス・パターンが 1 組のダイアゴナルモジュール群を示している。

【0082】

これら図 9 ~ 図 14 で左上に括弧書きで示した数字は予約済み入力ポート情報

の転送順序（列番号）を示しており、左端に示した数字は予約済み出力ポート情報の転送順序（行番号）を示している。それぞれのマトリックスの下に括弧で示した数字は、ダイアゴナル・サービス・パタンの種類を通し番号で表わしたものであり、同一番号は同一のパタンを示している。実施例で説明したダイアゴナルモジュール群は、この番号（１）～（４）に示す４つのダイアゴナル・サービス・パタンに対応していることになる。

【 0 0 8 3 】

なお、以上説明した実施例では説明を簡略に行うために 4×4 の比較的簡単なマトリックスを例に挙げて説明したが、マトリックスのサイズはこれに限定されないことは当然である。

【 0 0 8 4 】

【発明の効果】

以上説明したように請求項 1 記載の発明によれば、パケットの送られてくる入力ポートとこれらのパケットの出力先としての出力ポートを M 個ずつに分けてなる $M \times M$ のマトリックス状に配置されたグループにそれぞれのパケットの入出力の関係の情報としての転送要求の有無を表わした情報を割り振っておき、これらのグループについては、予め全入力ポートと全出力ポートを重複しないように等しい数ずつ選択してなる入力ポートと出力ポートの組み合わせとしているので、マトリックスの行方向にも列方向にもパイプライン処理が可能であり、パケットの転送処理の高速化を可能にすることができ、ポートの数の増大に十分対応することができる。

【 0 0 8 5 】

また請求項 2 記載の発明によれば、入力ポートと固定長セルの出力先としての出力ポートを M 個ずつに分けてなる $M \times M$ のマトリックス状に配置されたグループにそれぞれの固定長セルの入出力の関係の情報としての転送要求の有無を表わした情報を割り振っておき、これらのグループについては、予め全入力ポートと全出力ポートを重複しないように等しい数ずつ選択してなる入力ポートと出力ポートの組み合わせとしているので、マトリックスの行方向にも列方向にもパイプライン処理が可能であり、固定長セルの転送処理の高速化を可能にすることがで

き、ポートの数の増大に十分対応することができる。しかも、パケットを固定長セルに分解して処理するので、可変長のパケットについても高速処理が可能になる。

【0086】

更に請求項3記載の発明によれば、マトリックスの行方向にも列方向にも処理を行う本発明で全体としての処理が規則正しく進行し、処理の効率化を図ることができる。

【0087】

また請求項4記載の発明によれば、 $M \times M$ のグループに対して M 個からなるダイアゴナル・サービス・パターンを M 通り用意することで、いずれのタイムスロットでも1組のダイアゴナル・サービス・パターンによる予約処理が開始するので、効率的な予約処理が可能になる。また、 $M \times M$ 個の全グループがダイアゴナル・サービス・パターンの構成となるので、予約処理の優先度が結果的に均等に割り振られることになり、入力ポートによって優劣が生じるといった問題を生じさせないという効果もある。

【0088】

更に請求項5記載の発明によれば、予め定めた未来の所定のタイムスロットに対して、ダイアゴナル・サービス・パターンに該当するグループを起点として予約を行うことにしているので、これら未来のタイムスロットを重複しないように割り当てることで、効率的な処理を実現することができる。また、これらのグループを必要な個数連結して順次予約処理を実行させることで、スケラブルなスケジューラを実現することができる。

【図面の簡単な説明】

【図1】

本発明の一実施例におけるパケット交換装置の入力バッファ型スイッチのスイッチ部およびその周辺の構成を表わしたブロック図である。

【図2】

本実施例における M^2 個のスケジューリング・モジュールの処理動作を示した説明図である。

【図 3】

本実施例における V O Q の 1 つを拡大して図解した説明図である。

【図 4】

本実施例で 16×16 のクロスポイントスイッチにおけるスケジューリング・モジュールを表わした説明図である。

【図 5】

本実施例で第 1 のスケジューリング・モジュールのマトリックスの内部の信号配置を一例として示した説明図である。

【図 6】

本実施例で各スケジューリング・モジュールのスケジューリングの様子を時間の変化と共に表わした最初の説明図である。

【図 7】

図 6 に引き続いて本実施例におけるスケジューリング・モジュールのスケジューリングの様子を時間の変化と共に表わした説明図である。

【図 8】

図 7 に引き続いて本実施例におけるスケジューリング・モジュールのスケジューリングの様子を時間の変化と共に表わした説明図である。

【図 9】

スケジューリング・モジュールが 4×4 で構成される場合のダイアゴナルモジュール群の選び方とモジュール間の接続の順序の第 1 の例を示した説明図である。

【図 10】

スケジューリング・モジュールが 4×4 で構成される場合のダイアゴナルモジュール群の選び方とモジュール間の接続の順序の第 2 の例を示した説明図である。

【図 11】

スケジューリング・モジュールが 4×4 で構成される場合のダイアゴナルモジュール群の選び方とモジュール間の接続の順序の第 3 の例を示した説明図である。

【図 1 2】

スケジューリング・モジュールが 4×4 で構成される場合のダイアゴナルモジュール群の選び方とモジュール間の接続の順序の第 4 の例を示した説明図である。

【図 1 3】

スケジューリング・モジュールが 4×4 で構成される場合のダイアゴナルモジュール群の選び方とモジュール間の接続の順序の第 5 の例を示した説明図である。

【図 1 4】

スケジューリング・モジュールが 4×4 で構成される場合のダイアゴナルモジュール群の選び方とモジュール間の接続の順序の第 6 の例を示した説明図である。

【図 1 5】

V O Q を使用した入力バッファ型スイッチの一般的な構成を表わしたブロック図である。

【図 1 6】

スケジューリングをパイプライン化した従来のスイッチの構成の一例を示した概略構成図である。

【図 1 7】

従来のパイプライン化処理に使用されるパケットスイッチスケジューラのブロック図である。

【図 1 8】

従来のパイプライン化処理における接続許可情報の決定の様子を示した説明図である。

【図 1 9】

従来のパイプライン化処理における各インプットモジュールの予約順序を示した説明図である。

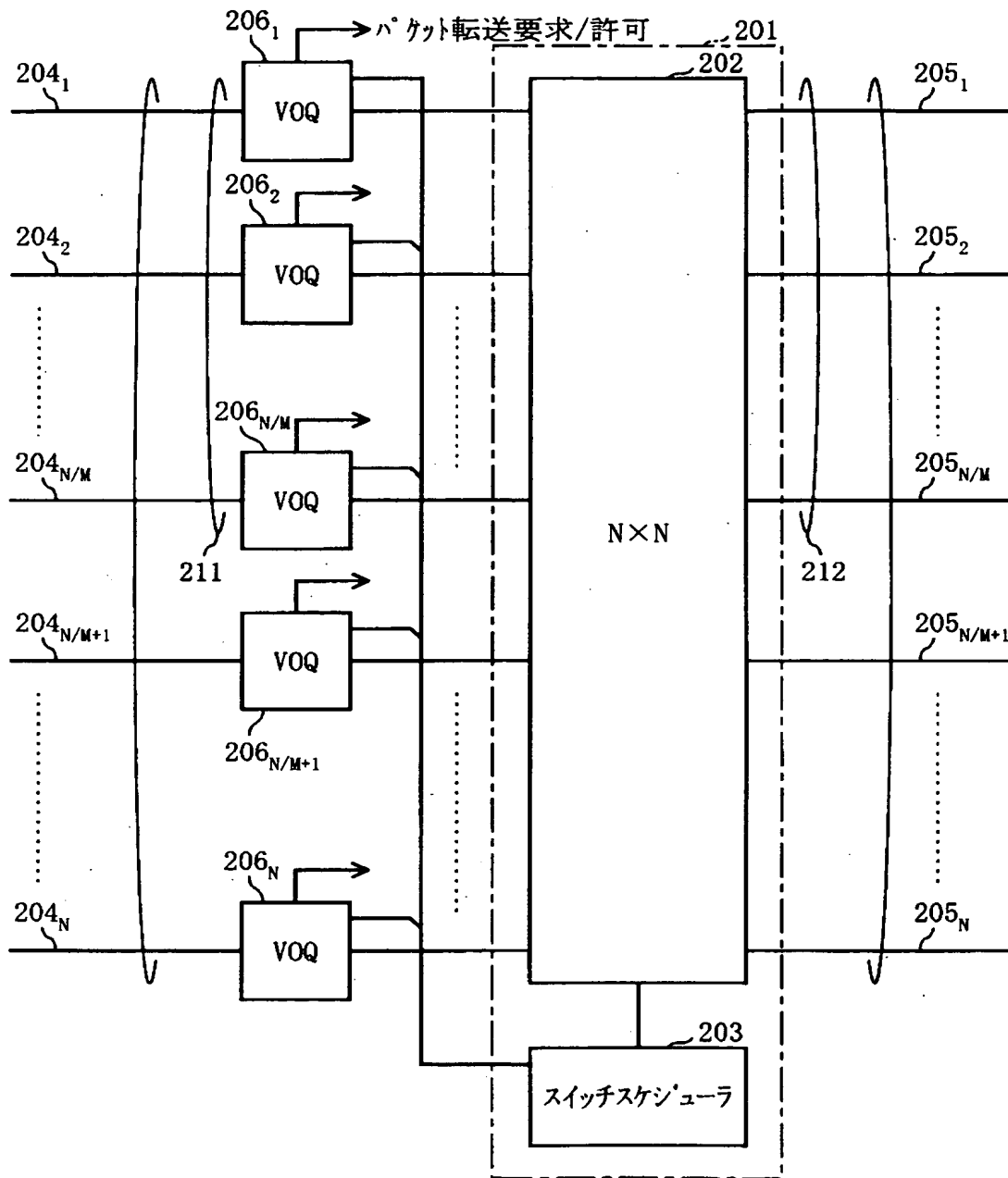
【符号の説明】

105 パケット生成分解部 (PAD)

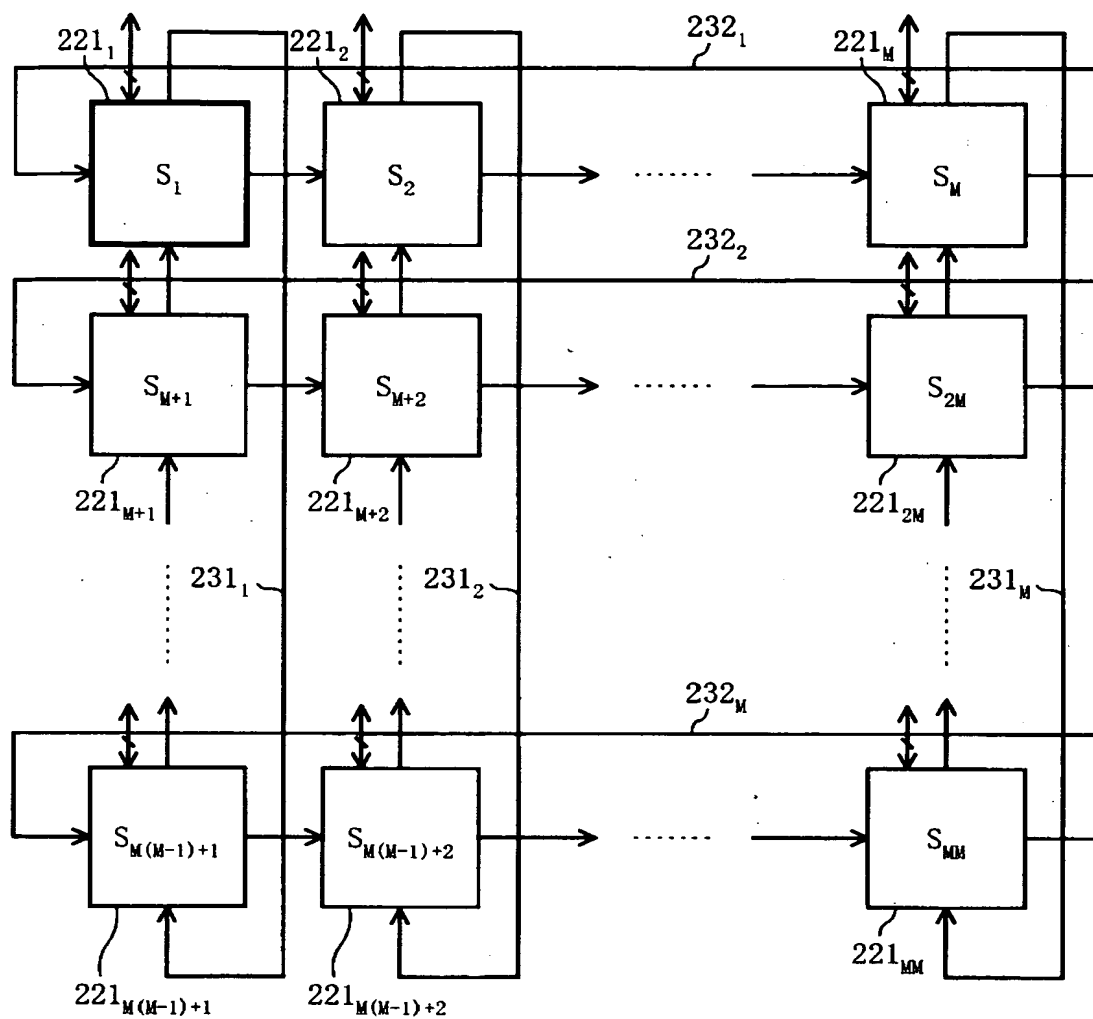
- 201 スイッチ部
- 202 クロスポイントスイッチ
- 203 スイッチスケジューラ
- 204 入力ライン
- 205 出力ライン
- 206 VOQ
- 221 スケジューリング・モジュール
- 231、261～264 予約済み出力ポート情報
- 232、265～268 予約済み入力ポート情報
- 241 固定長セル（パケット）
- 242 論理キュー
- 244 グループ

【書類名】 図面

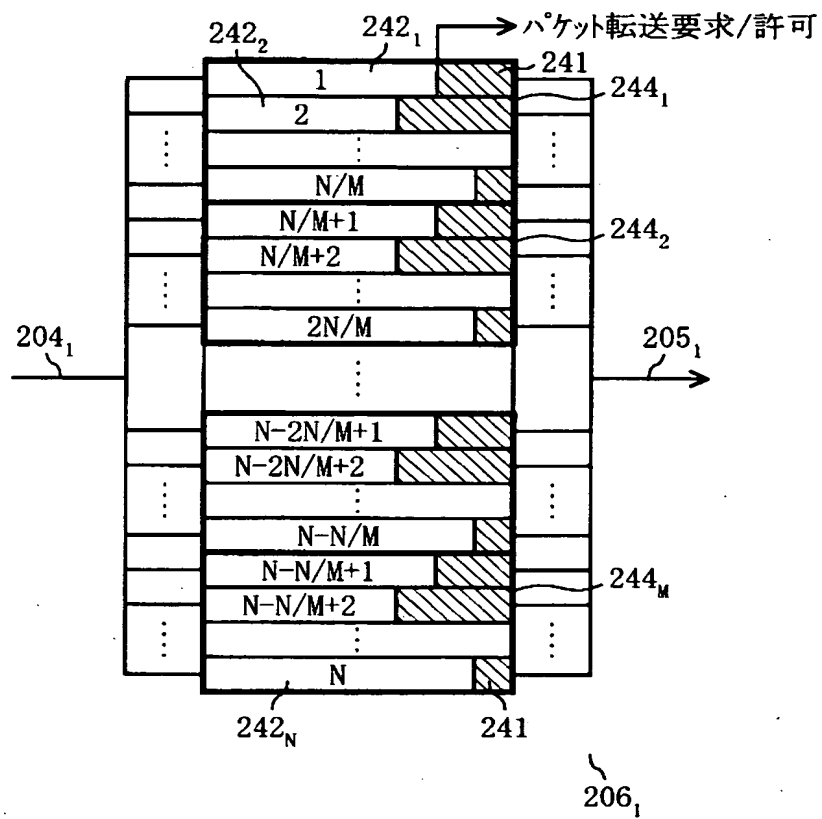
【図 1】



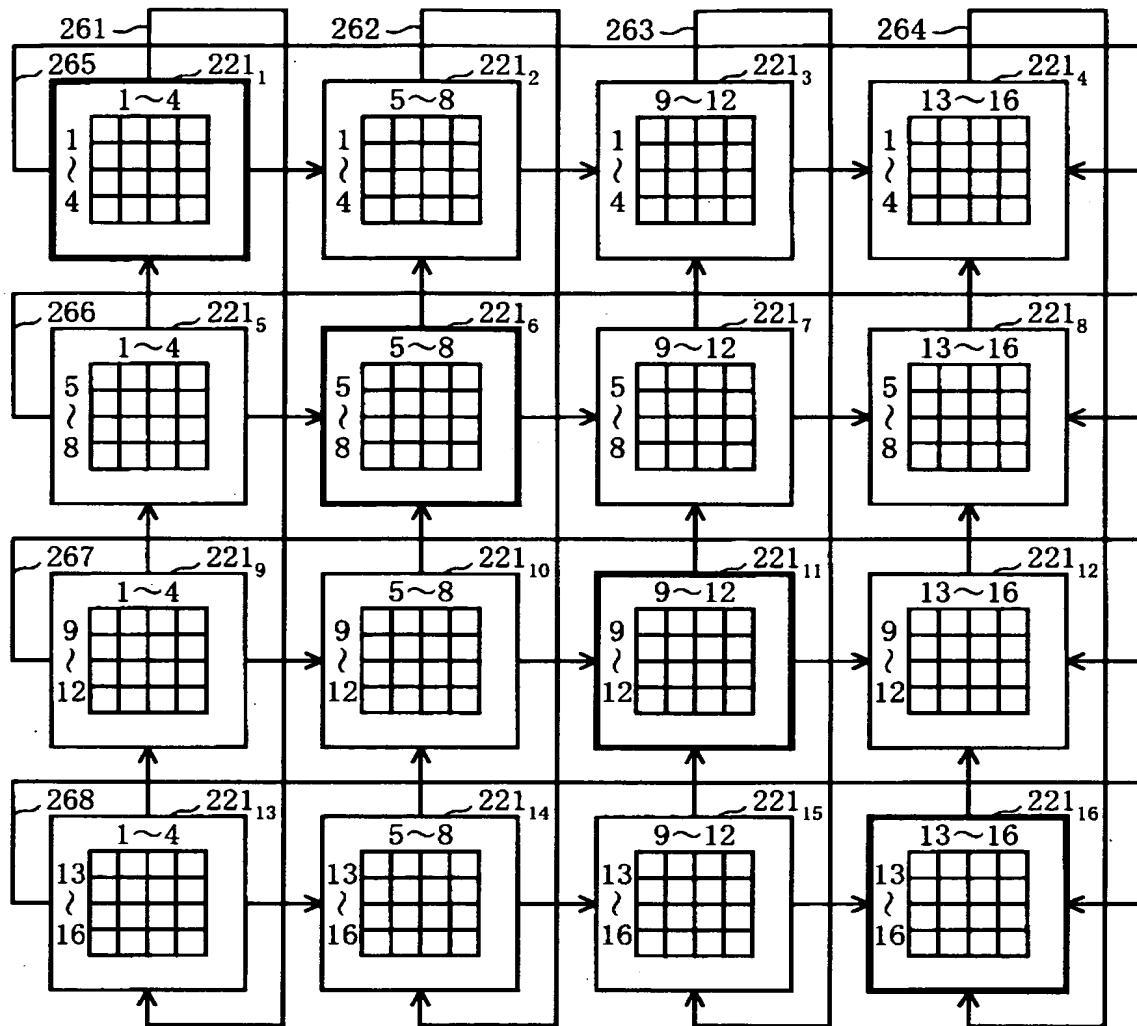
【図 2】



【図 3】



【図4】



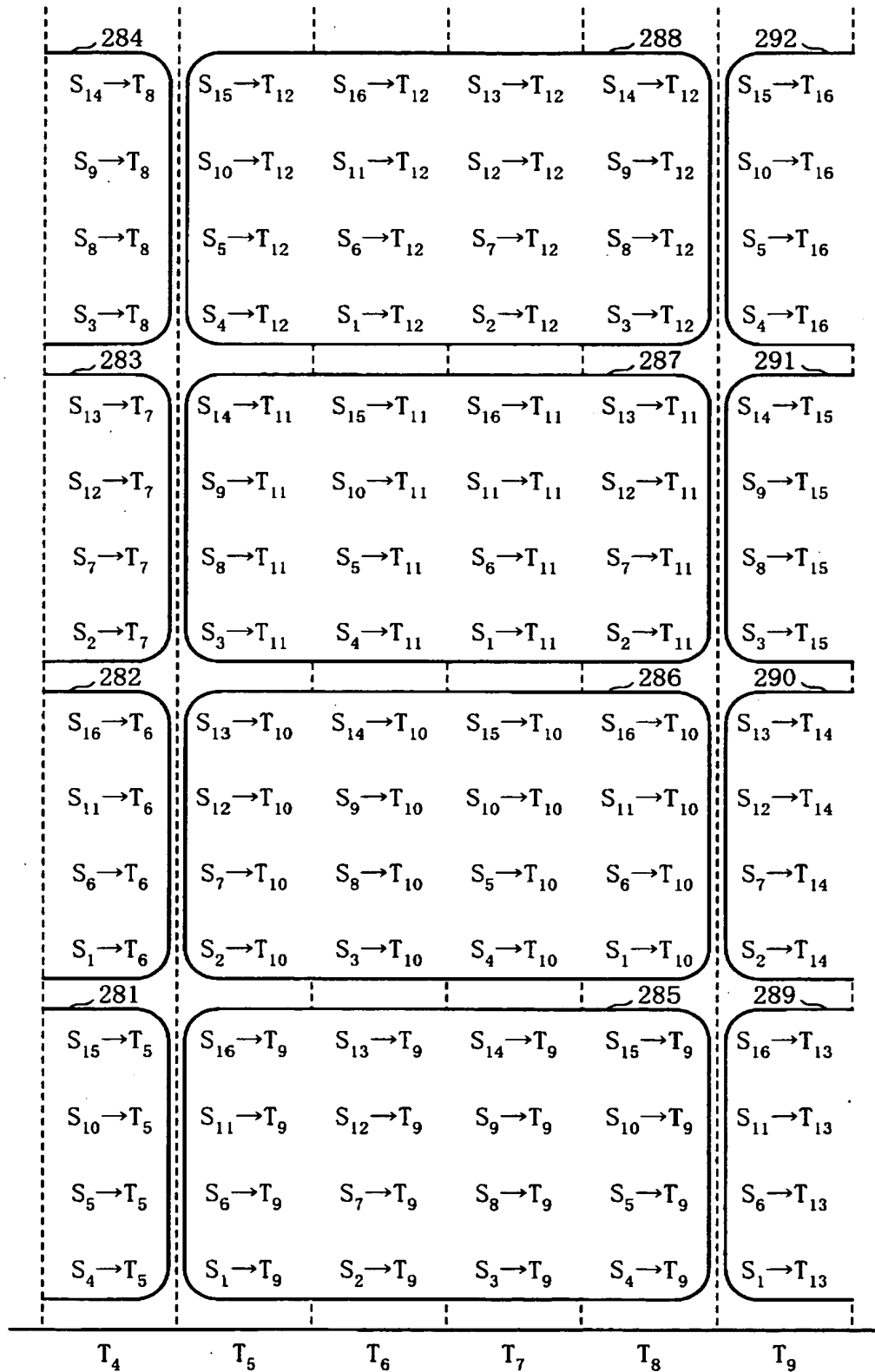
【図5】

	1	2	3	4
1	0	1	0	1
2	0	0	0	1
3	1	0	1	0
4	0	0	1	0

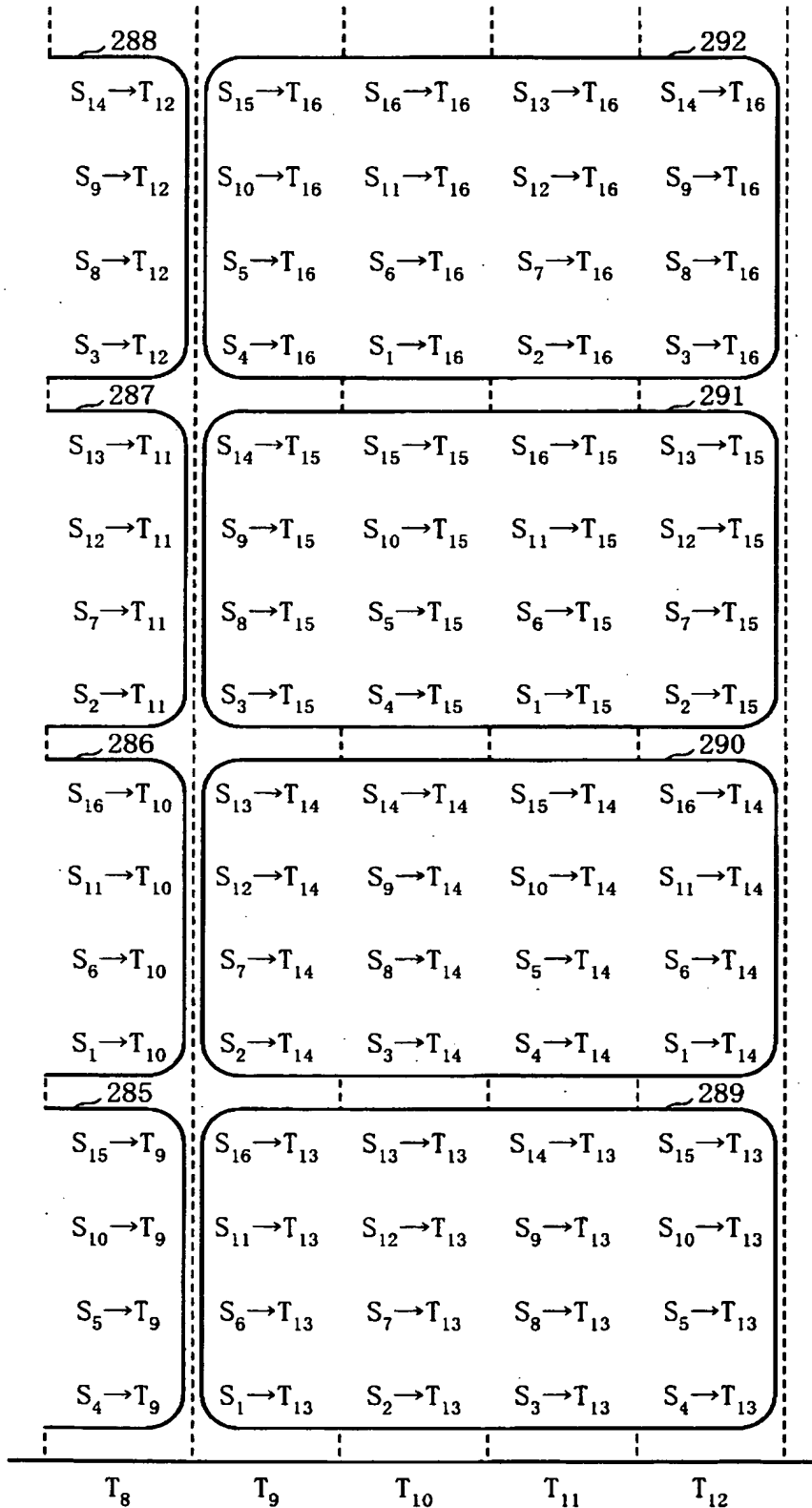
【図 6】

284				288
$S_{15} \rightarrow T_8$	$S_{16} \rightarrow T_8$	$S_{13} \rightarrow T_8$	$S_{14} \rightarrow T_8$	$S_{15} \rightarrow T_{12}$
$S_{10} \rightarrow T_8$	$S_{11} \rightarrow T_8$	$S_{12} \rightarrow T_8$	$S_9 \rightarrow T_8$	$S_{10} \rightarrow T_{12}$
$S_5 \rightarrow T_8$	$S_6 \rightarrow T_8$	$S_7 \rightarrow T_8$	$S_8 \rightarrow T_8$	$S_5 \rightarrow T_{12}$
$S_4 \rightarrow T_8$	$S_1 \rightarrow T_8$	$S_2 \rightarrow T_8$	$S_3 \rightarrow T_8$	$S_4 \rightarrow T_{12}$
283				287
$S_{14} \rightarrow T_7$	$S_{15} \rightarrow T_7$	$S_{16} \rightarrow T_7$	$S_{13} \rightarrow T_7$	$S_{14} \rightarrow T_{11}$
$S_9 \rightarrow T_7$	$S_{10} \rightarrow T_7$	$S_{11} \rightarrow T_7$	$S_{12} \rightarrow T_7$	$S_9 \rightarrow T_{11}$
$S_8 \rightarrow T_7$	$S_5 \rightarrow T_7$	$S_6 \rightarrow T_7$	$S_7 \rightarrow T_7$	$S_8 \rightarrow T_{11}$
$S_3 \rightarrow T_7$	$S_4 \rightarrow T_7$	$S_1 \rightarrow T_7$	$S_2 \rightarrow T_7$	$S_3 \rightarrow T_{11}$
282				286
$S_{13} \rightarrow T_6$	$S_{14} \rightarrow T_6$	$S_{15} \rightarrow T_6$	$S_{16} \rightarrow T_6$	$S_{13} \rightarrow T_{10}$
$S_{12} \rightarrow T_6$	$S_9 \rightarrow T_6$	$S_{10} \rightarrow T_6$	$S_{11} \rightarrow T_6$	$S_{12} \rightarrow T_{10}$
$S_7 \rightarrow T_6$	$S_8 \rightarrow T_6$	$S_5 \rightarrow T_6$	$S_6 \rightarrow T_6$	$S_7 \rightarrow T_{10}$
$S_2 \rightarrow T_6$	$S_3 \rightarrow T_6$	$S_4 \rightarrow T_6$	$S_1 \rightarrow T_6$	$S_2 \rightarrow T_{10}$
281				285
$S_{16} \rightarrow T_5$	$S_{13} \rightarrow T_5$	$S_{14} \rightarrow T_5$	$S_{15} \rightarrow T_5$	$S_{16} \rightarrow T_9$
$S_{11} \rightarrow T_5$	$S_{12} \rightarrow T_5$	$S_9 \rightarrow T_5$	$S_{10} \rightarrow T_5$	$S_{11} \rightarrow T_9$
$S_6 \rightarrow T_5$	$S_7 \rightarrow T_5$	$S_8 \rightarrow T_5$	$S_5 \rightarrow T_5$	$S_6 \rightarrow T_9$
$S_1 \rightarrow T_5$	$S_2 \rightarrow T_5$	$S_3 \rightarrow T_5$	$S_4 \rightarrow T_5$	$S_1 \rightarrow T_9$
T_1	T_2	T_3	T_4	T_5

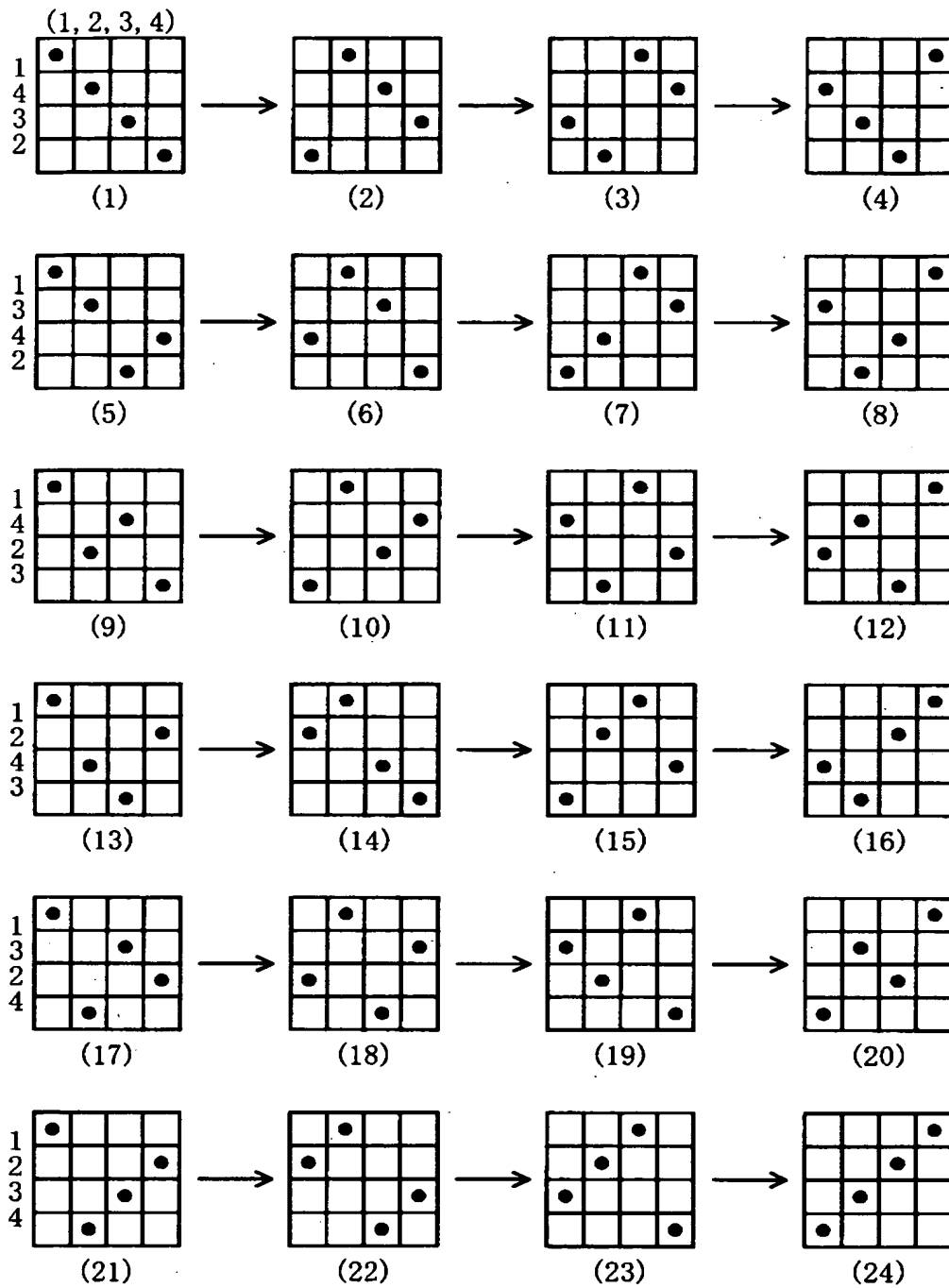
【図 7】



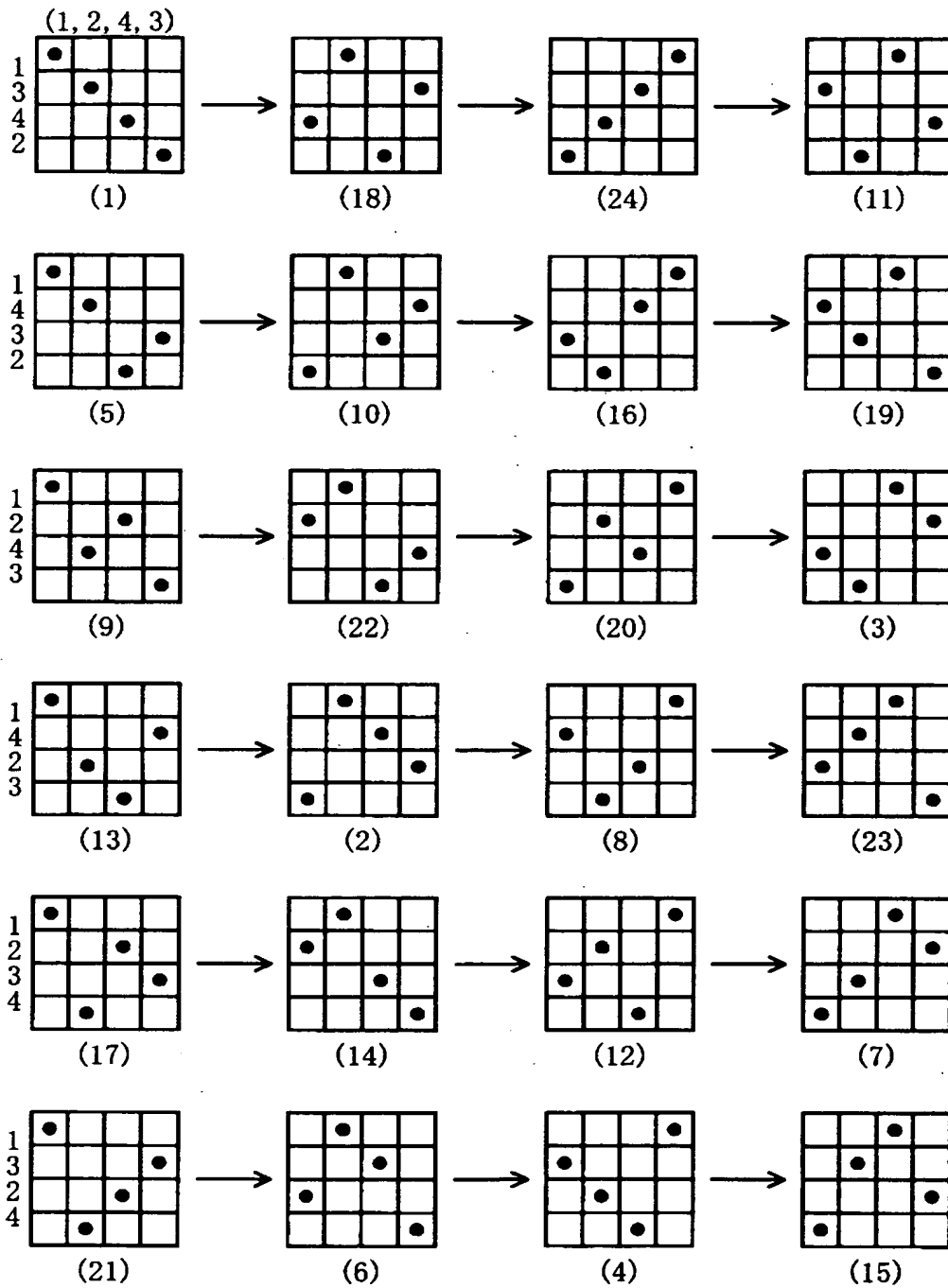
【図 8】



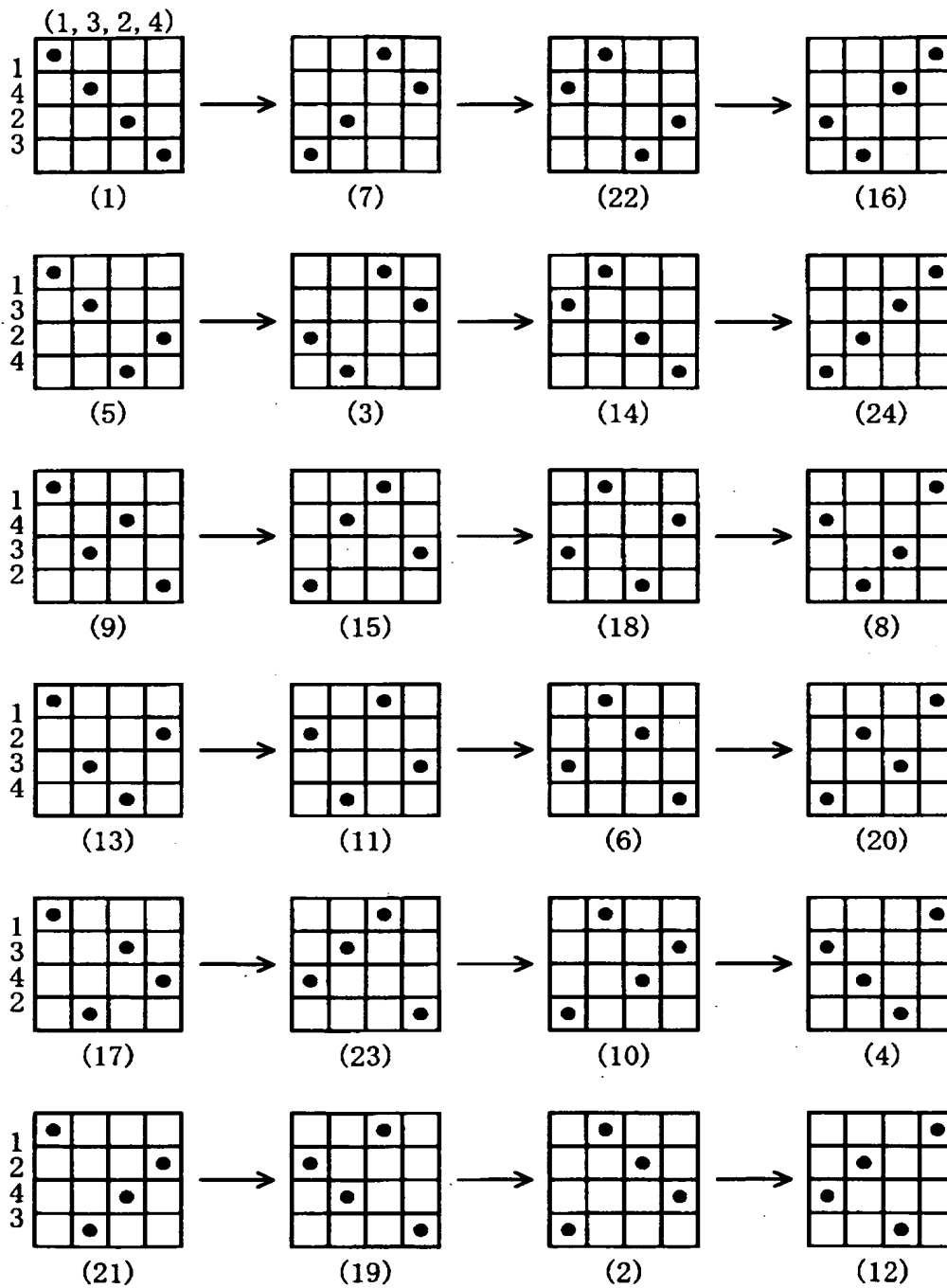
【図9】



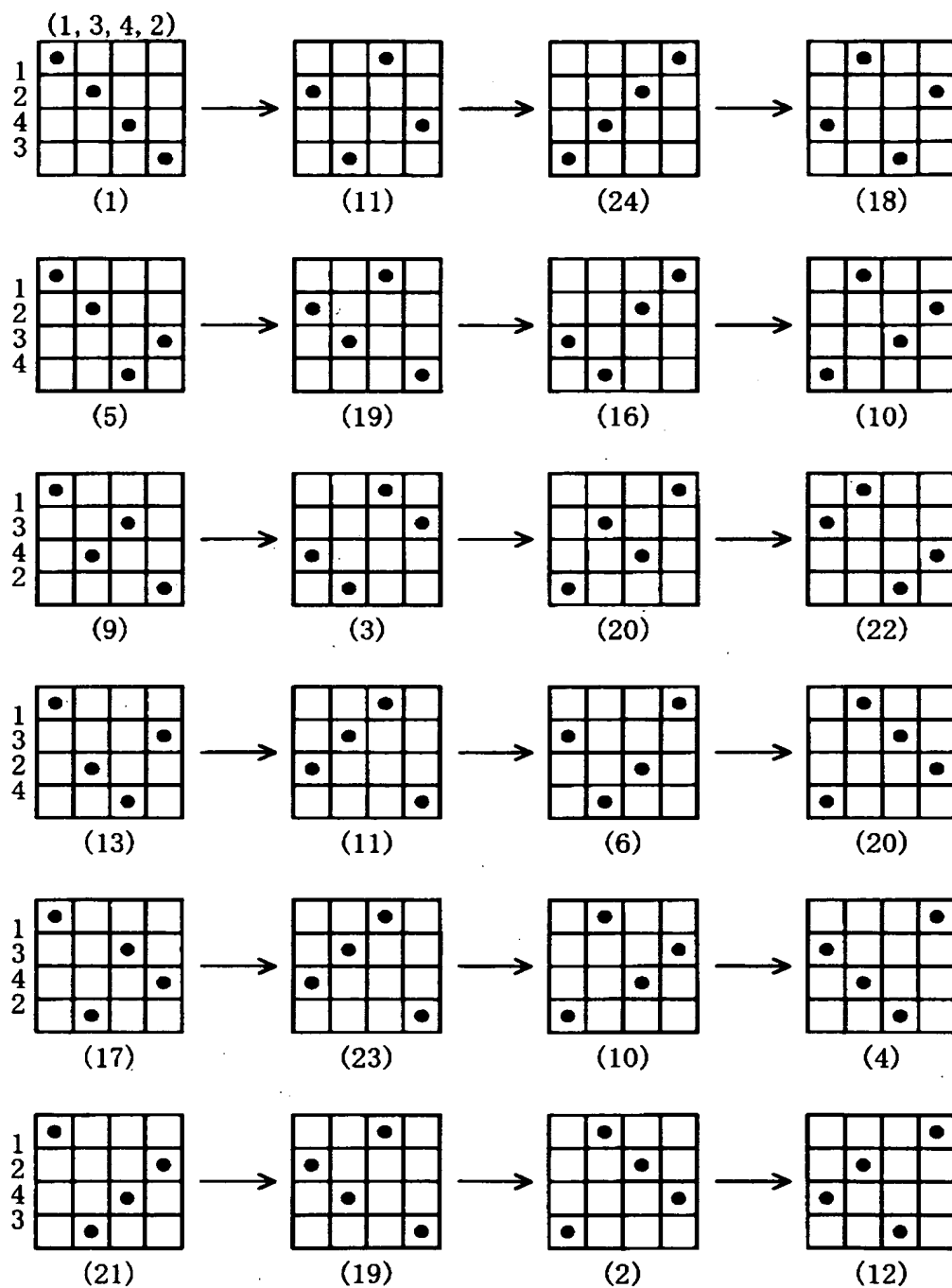
【図 1 0】



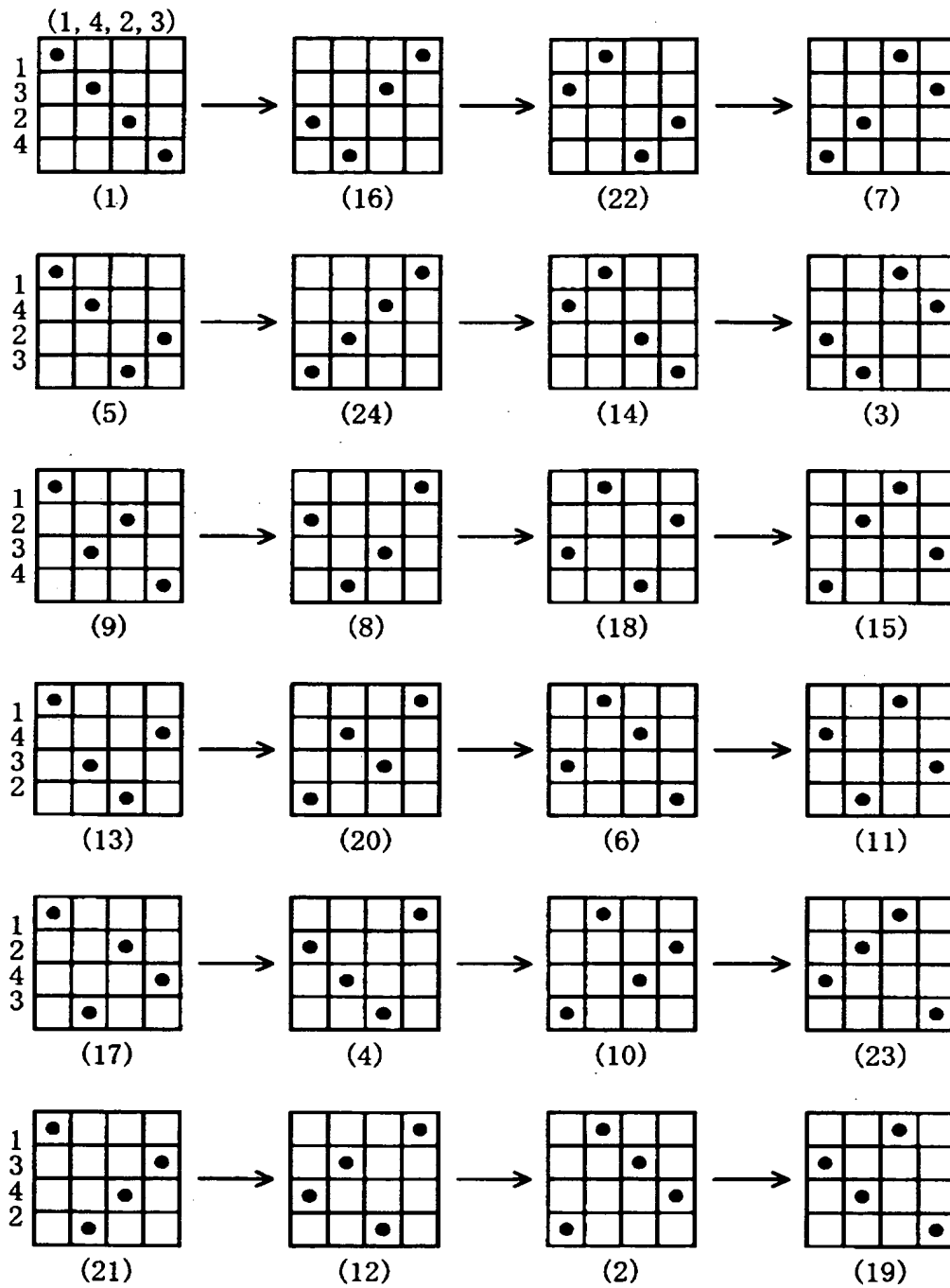
【図 1 1】



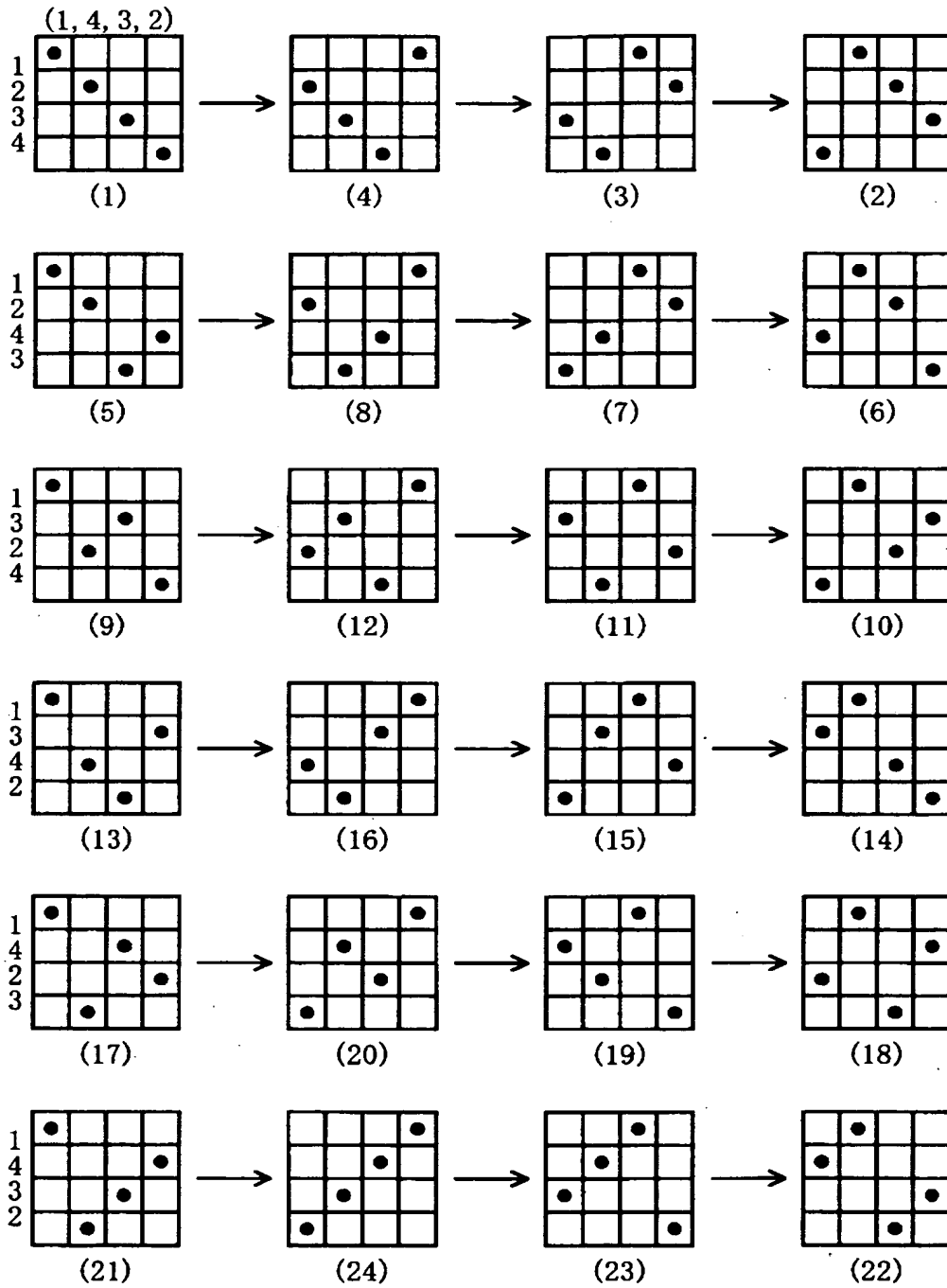
【図 1 2】



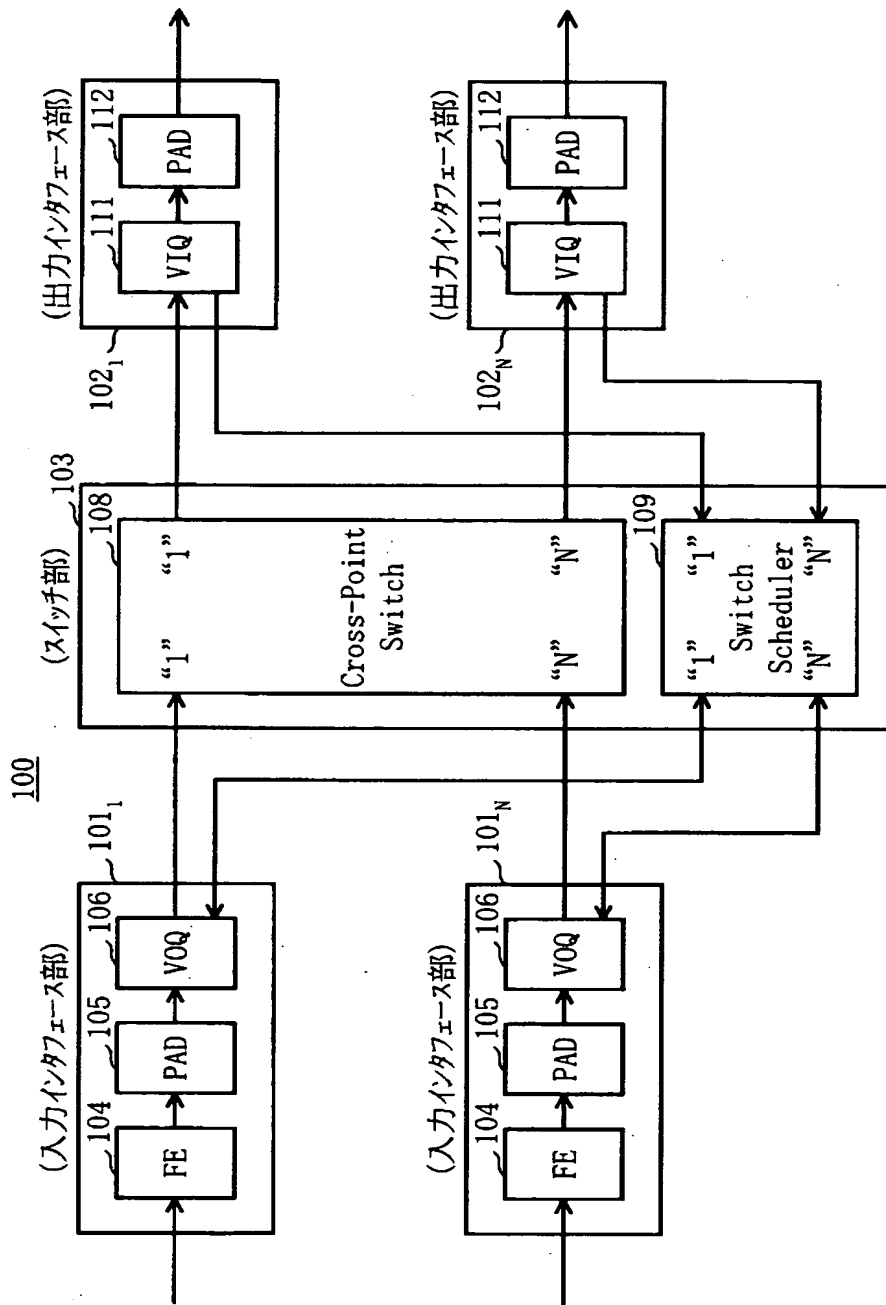
【図13】



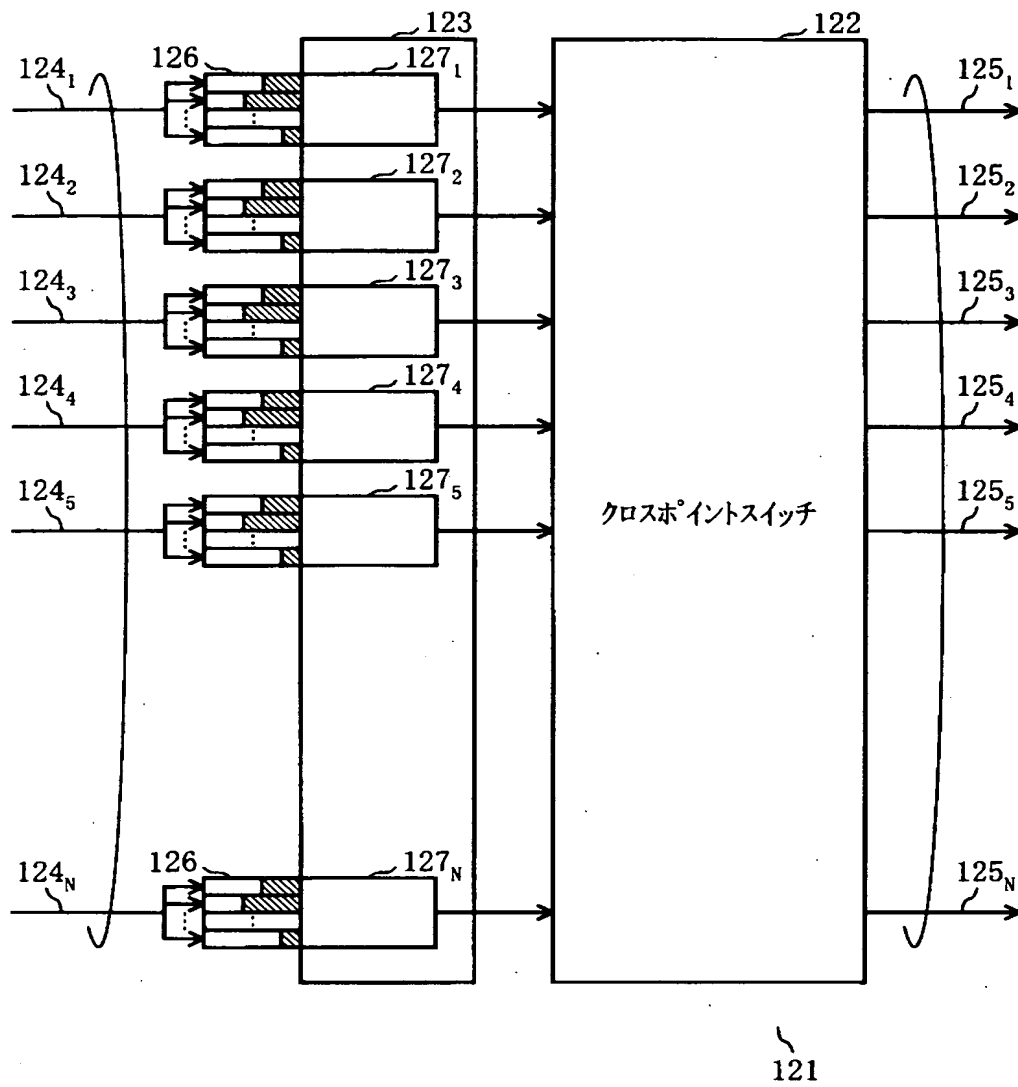
【図 1 4】



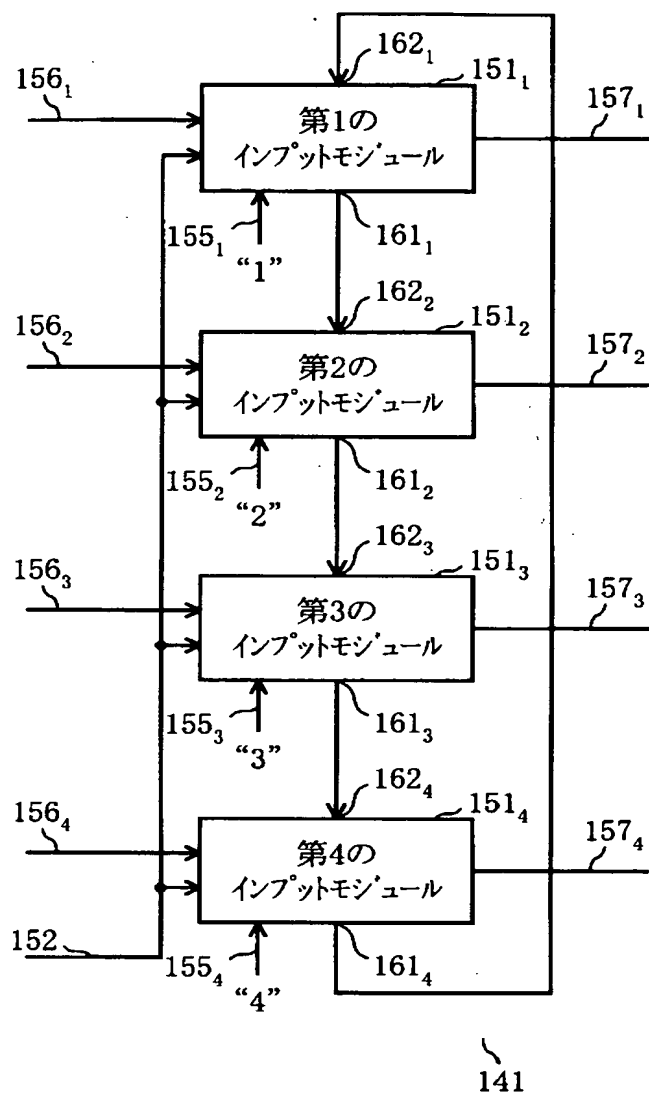
【図15】



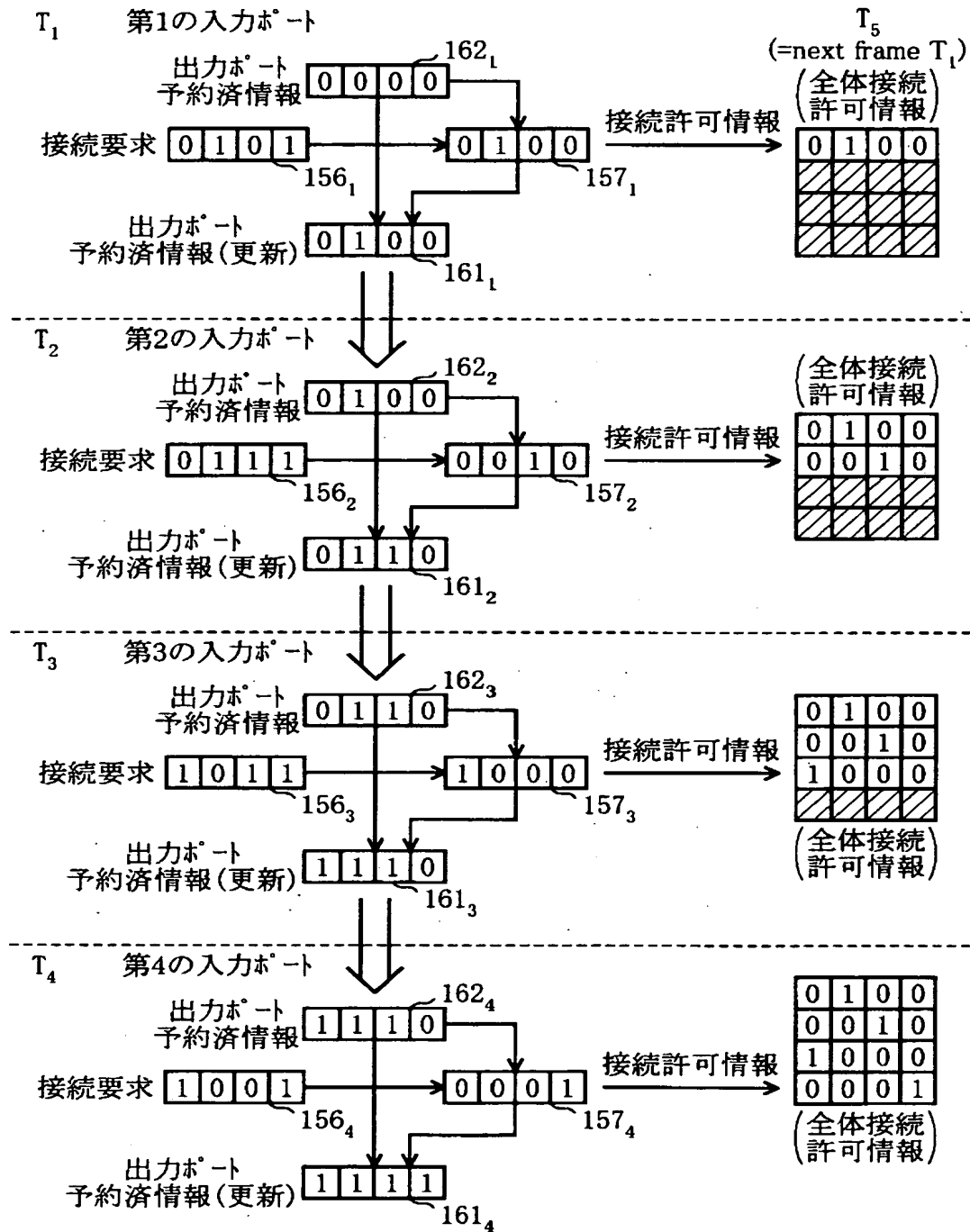
【図 1 6】



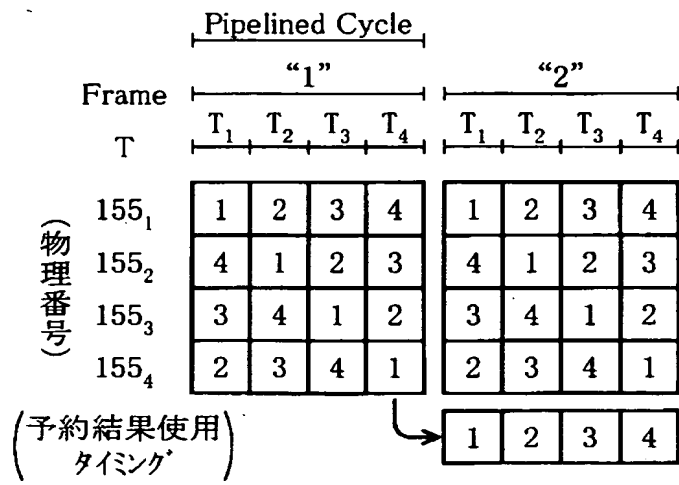
【図 1 7】



【図 1 8】



【図 1 9】



【書類名】 要約書

【要約】

【課題】 入力ポート側のパイプライン化処理を実現した場合よりも更に高速処理が可能なパケット交換装置を得ること。

【解決手段】 各入力ライン 2 0 4 から入力されるパケットは固定長セルに分解された後、対応する論理キューとしての V O Q 2 0 6 に入力されて出力ライン 2 0 5 ごとに振り分けて格納される。これらの格納情報はスイッチスケジューラ 2 0 3 に入力される。スイッチスケジューラ 2 0 3 は、クロスポイントスイッチを構成するマトリックスの数だけのスケジューリング・モジュールを備えており、入力ライン 2 0 4 だけでなく出力ライン 2 0 5 についてもパイプライン化した予約処理を行い、予約結果に応じて各タイムスロットごとに固定長セルの入出力間のスイッチングを行わせる。出力ライン側のパイプライン処理により、高速処理が可能になる。

【選択図】 図 4

認定・付加情報

特許出願の番号	特願2000-302551	
受付番号	50001276600	
書類名	特許願	
担当官	第八担当上席	0097
作成日	平成12年10月	3日

<認定情報・付加情報>

【提出日】	平成12年10月 2日
-------	-------------

出 願 人 履 歴 情 報

識別番号 [000004237]

1. 変更年月日 1990年 8月29日

[変更理由] 新規登録

住 所 東京都港区芝五丁目7番1号

氏 名 日本電気株式会社